

GenCore version 5.1.4-p5,4578  
Copyright (c) 1993 - 2003 Compugen Ltd.

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

## ALIGNMENTS

RESULT 1

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

DEFINITION

ACCESSION

VERSION

SOURCE

DEFINITION</p

JOURNAL	Curr. Genet. 12 (3), 175-184 (1987)
MEDLINE	88210533
PUBMED	2835179
FEATURES	Location/Qualifiers
source	1 .4193
repeat_region	/organism="Kluyveromyces marxianus" '/strain="ATCC 12424" '/ab_xref="Taxon:4911" '/clone_lib="PHCG3" 368. .276
repeat_region	/note="direct repeat 1" 307. .315
promoter	/note="direct repeat 1" 316. .320
repeat_region	/note="TATA-Box like sequence" 342. .355
promoter	/note="TATA-Box like sequence" 400. .408
repeat_region	/note="direct repeat 2" 406. .412
repeat_region	/note="TATA-Box like sequence" 416. .424
misc_feature	/note="direct repeat 2" 459
misc_feature	/note="pot. transcription start site" 482. .483
misc_feature	/note="pot. alternate transcription start site" 612. .3149
CD5	/codon_start=1
	/protein_id="CA29353.1"
	/ab_xref="SWISS-PROT:P07337"
	/translation="MARDVWQSLSLNQDKLISLAVDFWHTKKIERLGPAPRVS DCPQNGITRGTKFDVGVEGQNLQGTLASTDILETAGKLMAKSTIAKNAAVLIG TMMQRGIGGFGFSESEDPLAGMATSVNGMOCGEGTATAVKHFVCDLLEDRESS NSLIVSEALRELYLEFPLRALKHANCVIMWAKYNGDICSQSKLILDRBWK DNLMSDWFGVYTTAAIKNQDILIEPGPWRTRALVSILNSREQIT"EDDVDRVR QLMKIKFVWNLKINGIENGPESTSNNTKTSOLRELAQDSTVLLNNKNNLTSK ERKQYHVTGPAKATVSSGGASMSMYYVSPYIUVNKGKEDTYVAGYSRKG GLAESSLLIAKPADENAGSIKAKSYNSPYVYERS2DEEPHVTNKSINYVHPSKHE KVDPKNPYEFVLTGQYVQPDGDXFISLQYQGSCLFYLNEDLIDOKINQERGSFC GAGTKEKRRKLTQKQVYNNRVEYSGPQSLVGRGAGFQVAKDDEBIRN AELAAKHKDVKAVLILGNGENETDRENNDPKRTNEVLRVAKNPTVNTDQ TEVEPMPLEENALNQAWMGNEGNAIAYLVGQWPNKLSLWPKLQDNDQFLN FKEFGRVWYSEDDIVFYRKEVTLQRKVAFFEGYGSYTFEFLDISDKYTTDDKIDIS WVVKNTQDKGSEDEVYFVAKLNSKVRKPEKTYLPERGEKKTINIEELKD ATSYFNEELGWVHWRAGEYLIVSUGTSSDDILSYKPERKVKEDLYWKLG"
repeat_unit	3314. .3322
misc_feature	/note="inverted repeat A" 3442. .3445
misc_feature	/note="yeast termination consensus sequence" 3472. .3474
misc_feature	/note="yeast termination consensus sequence" 3498. .3501
repeat_unit	/note="yeast termination consensus sequence" 3541. .3543
repeat_unit	/note="yeast termination consensus sequence" 3652. .3652
repeat_unit	/note="inverted repeat B" 3655. .3666
repeat_unit	/note="inverted repeat B" 3714. .3722
BASE COUNT	1338 a 675 c 941 g 1239 t
ORIGIN	
Query Match	17.1% Score 427; DB 8; Length 4193;
Best Local Similarity	50.7%; Pred. No. 1 5e-42;
Matches	1253; Conservative 0; Mismatches 1155; Indels 63; Gaps 7;
QY	1 AGGGCTGATATGATGCTGAGGCCACCTTGAGAAGCTCCTGCCGAGAACCTCGAT 60
Db	612 ATGCTCAAATTTGAGTTGAGCTAACAGTTATGAGTAAATTGACAGGATGAAAGATTC 671
QY	61 CTGCTGCGCTGATGACTCTGGACACAAAGGCTCTCCCAAGCATGGAGTCCCCT 120
Db	672 TTACTTCTCGCAGTGTATCTTGCCATAATGGATTAAGAGATTAAGAGTTGGAATTCAGCG 731
QY	121 CTCCGCTTACAGANGGCCAACAGGGCTGGTACACTTCAACAAACTCGCTCG 180
Db	732 GTGAGGTTCTGTAGTGGCTCAAATGGATTAAGAGATTAAGAGTTGGAATTCAGCG 791
QY	181 GCGGCTGCTCCCTGAGGCAACGGCGTGGTCCACLTCAACAAACTCGCTCG 300
Db	852 ACAGGAGTAGTGTGCTAAGAATGATGATGCTGAGAAATGCTGCTGATTTGGT 911
QY	301 CCGACTATCACATGCAAGCTCCCTCTGGTGGCATCTACTTTGATCGGATCTCG 360
Db	912 CCAACACAAACATGCAACTGGCTCTGGTGGTGGCATCTCG 971
QY	361 GATCGGTTCTGGGGCTGGAGCTGGCTCATCGCGCATTCAGACACTGGA 420
Db	972 GATCCATATCTGCTGTAAGCTGACTTCTCTGTTAAGTGTAAAGTGTATGCGAAGG 1031
QY	421 GTGCAGCTACGATGAAAGACTTTTGCTGCAATGATCAGGAGCACGGCATGTTG 480
Db	1032 ATGGCTGACGTTAACGATTGTTAACCTGGAAGACCGTTCTCTG 1091
QY	481 CAGAGCTGTCACGGCGCTCCGAACTAGCCTCCGGTCCAGATGGCT 540
Db	1092 AACTCATTTCTGCAAGAGGAATTACTTGGAGCCCTCAGATGGCA 1151
QY	541 GTGCCAGACGCCACCGGGTGCCTCAGACGCGCTACATGCAATSGCTG 600
Db	1152 GTTAACATGCAATCTGTTGATATGACTCTTACAGGTCATGGATCATGGATCAT 1211
QY	601 TGCAGGGACCCCTAAATCTGATGGATGCTTCGAAGGATGGTGGATGGC 660
Db	1212 TGCCTCCATTCACGAAAGCTATGACGATTTGAGGACAGTGGACGGT 1271
QY	661 CTAACTGAGGACTGGAGGACATACAGTACACAGAAGCCGTGGAGCT 720
Db	1272 ATGTTATGCTCGACTGGTGGTACATATACGCTGGCACTATCAAGATGGTG 1331
QY	721 GACCTGAGGAGCCGGACTCCACGCTTCAGCTGGAGGAAACACTCAAGTCAACGT 776
Db	1332 GATATGGAGTCTCTGGACACAGAGTGGAGACCGCTTGTAGTCTCTC 1391
QY	777 -CTCCAAAGGAAGCCCTTATCCGCGATGACCGAGGCTTGGAGGT-TCAG 834
Db	1392 AACTCAGAGAACAAATCTACTGAGATGGTGTGATGCTGTTAGCAAGTC 1451
QY	835 TTGGCTCAAGAGTGTGCTCGCTGGAGTAC 882
Db	1452 ATGATTAAGTTGTTGTCACAAATTTAGAGAACACGGTATGTTGGAGAATGCCAGA 1511
QY	883 ACGACTGTCACACACCCCGAACGGAGCTCCCGGAGGTGGCAGAGGGC 942
Db	1512 TCTATTCACACACGACCGAACGGACCTGGACCTGTTAGAGGAGATGCTGCT 1571
QY	943 ATCGCTGCTGTTGAGAACAGAACAGTCTGCCCTTGAGCAGAAGAGAACCTG 1002
Db	1572 ATGGTTTATGAGAACAAATATCTTACCTTAAGAAGAACATATCAT 1631
QY	1003 ATGGGGCCCAAGGCCAACGGCAATACCGGGACCTGGGACTCTGGACTCTG 1062
Db	1632 GTCATGGCCAATGCTAAAGAACAGTCTGCCCTGGTGTGATCATCTATC 1691
QY	1063 GCCTACTGCGAGTCACCTCTTGAGCCTCTGAGCAAGCAGTCAGACGCCATCG 1122
Db	1692 TCCACTATGTTCTCTCGTGTAGGATTCATAGCTGCAAGAGGTCG 1751

QY	1123 TACACCGTGGCCTCTA-CACCAACCGTCTCCCATCTAGCGGAGCGTCCTACGC 1180	QY	2167 AGCGTGTGCCCTCCCTGAGAACACCGGTCGCGGCCGCGCACGGGGCCGCC 2226
Db	1752 TACACCGTGGCCTCTA-CACCAACCGTCTAGCGTGGAGGTTGCGAGAGTAGTTGATC 1811	Db	2823 GATATTTCAGTGTAGTGTAGATACTTGGTATAATTTCGCTCCGAGGTGTC 2882
QY	1181 CGCAGGGCCTCCTGGCATGGCTGGAGGGCTCTCAAGA-GCCCTCTGGTACCCCTAA 1238	QY	2227 TAGTCAGCCCTCCAGGCCAGATTAACCGCCCTCAAGGAGCTCAAGGCTC 2286
Db	1812 GATCTGCAAACACAGCAGATGCTGAAATGCTGGATAATGCCGATTAATCTCCAA 1871	Db	2883 GTCATCTCAGGCCCTAACTCTAAGGTCTGAGACCGGTTAAGAGTCAGGATC 2942
QY	1239 CCGCAGCACA-.....-TGTAGACTCTTCTTCAAGAGCAG 1275	QY	2287 GCAAGGTCGAACTTCGACGCCGGAGACAGGAGCTGACATGGAGCAAGAG 2346
Db	1872 CGCAGTAGAGAGAGATCTGAGGTGAAGAACCATTCACGTTACCAAAAGTCATAGTCC 1931	Db	2943 GAAAGTTCTATTGGACACGGTGAAGAACAGTATATGACTGAAATGAA 3002
QY	1276 GACATGACCTGGGAGACTACTACACCCANGGGCAGACG-.....-TGGTAC 1326	QY	2347 TAGTCGCTGGTATTTGAGGAGGGGATCAGTGGTGTGAGAAAGGGGACTAT 2406
Db	1932 AATGTTCACTTATTTGATTTCAACATGAGAAAGTGGATCCAAAGAACCTTACTTT 1991	Db	3003 GATGCAATTCTACTTTAACGAAGAGCTCGTCTCGTAATGGACGCTGAG 3062
QY	1327 GCCACATGGAGGCGACGTAACACGCCAGGAGACTCACCCTACGAGCTCGGCC 1386	QY	2407 GAGTTATCGT 2417
Db	1992 GTCACCTAACCGTACGGACAGTACGGCTCAAGAAGATGTTGATATCTTCAGTC 2051	Db	3063 TTTGTTTCAGT 3073
QY	1387 GTCAGGGCAGCAGCAGGAAAGGGCAGACGAGCTGGTGTGAGAACGCCACAG 1446	RESULT 2	
Db	2052 GTTATAGGTCCTGTTGTCCTACTTACAGTATGAGTGTGATTTGACCAAGACAC 2111	DEFINITION	AF486581
QY	1447 CAGTCGCCGGCCTGCCCCTCTGGCCTGGCACCACGGAGAGGGGCCATCAAT 1506	ACCESSION	AF486581
Db	2112 CAGAAAGGGTGTAGTTCTGCTTGGAGCTGGTACCAAAAGAAACCAAAGTACT 2171	VERSION	AF486581.1
QY	1507 CTCTCAAGGCCACACGCTACAGTCAAGTCTGAGTCGGTCTCCGACCCACCTAC 1566	KEYWORDS	
Db	2172 TTGAGAAAGGGCAGACCATCTGTCGGCACGGCTCCGCTGGGGCTGCAAGTC 2231	SOURCE	uncultured bacterium:
QY	1567 CTCAAGGCCCTAACGGGACTGGTGTGAGTGTGAGTCGGTCTGCTCAACTCGT 1626	ORGANISM	Bacterium; environmental samples.
Db	2232 TTGTTGGGA-.....-ATTCGGTGTGAGGGATTCACGCTGGTCAATAGCG 2282	REFERENCE	1 (bases 1 to 38269)
QY	1627 ATTCACGACGGCCGAATCGAAAAGTCGTCGTCGCCCCTCGCAAGGAGCACCGAGTC 1686	AUTHORS	Cappellano, C.M., Courtous, S., Ball, M., Francou, F.-X., Normand, P., Heylinck, G., Guerneau, M., Martinez, A., Hopke, J., Kolvek, S., August, P.R. and Martínez, P. and Permodet, J.-L.
Db	2283 ATCGATGATGACCGAGGATTAACCGACGGAAATTAGCAAGCTAACGATGATAAGCT 2342	JOURNAL	Recombinant Environmental Libraries Provide Access to Microbial Diversity for Natural Products Drug Discovery
QY	1687 ATTCATCTGCGCGGCCCTAACGGGACTGGAGGACGGGGCCGACGGGGAGATG 1746	2 (bases 1 to 38269)	
Db	2343 GTGTTGATAATTGATGTTAATGTTGAAATGGAAACGGTATGACAGAGAACATG 2402	AUTHORS	August, P.R. and Martínez, A.
QY	1747 AAGCTCCGGCTGCTGGACCGACGGCTACGGGGCCGAAACCCAAACACC 1806	JOURNAL	Submitted (21-FEB-2002) Molecular Biodiversity, Aventis Cambridge Genomics Center, 26 Landsdowne Street, Cambridge, MA 02139, USA
Db	2403 GATTGCCAAAGAACAAATGATTAGTCTGCTGTGTTGAAAGGAATCCCAAATCT 2462	FEATURES	Location/Qualifiers
QY	1807 GTGCTGCGCATGGAGCGGGCACCCGGAGGAGTCGCTCGACGCCAGGGCGCC 1866	source	1. -38269
Db	2463 GTTATCGTTAACCAATCAGGTACCCCTGTTGAGTCCCTTGGTTAGAGAGGAAATCGC 2522	CDS	/organism="uncultured bacterium"
QY	1867 GTCTTCCAGGCCCTGCTGGCGCAACGAGACGGGCACTCTCAATGGACCTCTGCTT 1926		/db_xref="taxon:77133"
Db	2523 CTAGTCAGCTGGTGTAGGTTAATGAAATGGTATGCTGCTCTGGCAATTAAAGTGGAGATAAT 2582		/note="isolated from soil"
QY	1927 GGGCACTACACCCCTCGGGCAAGCTGCGCTCAAGCTGCCATCTGGCACTT 1986		/note="isolated from soil"
Db	2583 GGAGACGCTGGTCTCAATGGTAGTATGCTGCTCTGGCAATTAAAGTGGAGATAAT 2642		/note="confers resistance to kanamycin in Escherichia coli; similar to known cognate proteins"
QY	1987 CGCGGTTCTCACTTCGACGGGCCGGCACGCTGCTACGGGAGGACGCTAC 2046		/codon_start=1
Db	2643 CGACCTTTAAACTCAAGACGAGTGTGGAGAGTTACGGTGTGAGATCTT 2702		/transl_table=11
QY	2047 GTCCGGTACAGGACTTACAGGTTGCTGGCAACAGACGCTCAATTCCCTTGGCAAGGC 2106		/product="aminoacyl-lysine 3'-N-acetyltransferase"
Db	2703 GTTGGTATAGGFACTAGCAAGAGCTTCCTAAAGAAGGTAGCTTCCTCGTGGATAGTT 2762		/protein_id=AAH2007.1"
QY	2107 CTGTCCTACACCCTTGCTTCCATCTCCTCCGGTCTCAGAGGGCAAGCTG 2166		/db_xref="GI:1956966"
Db	2763 CTATCGTATAACACATGGAACCTAGATATTCTGACTCTCAAGGTAATGATAGATA 2822		/translation="MTRRATSSLADDLSAIGLADGLVIALRQWKTGIVGGPDAIDARVDVGPAGTIGYDWDLDRDPSMPHIAFDPERSRSDNGWPEALRTPLALRGSPGAGMAAGWMTADHGXGSPICKLVEKGKIVMLGAPDTMLLHADPFRNIRLVEAGDNGWWREREFSPDGLPDPYATIIVEALATGKKGKRGVGEASVLPAAANVAFGDWLWKGKTL"
BASE COUNT		ORIGIN	16.2%; Score 4.06.2; DB 1; Length 38269;
Query Match		Best Local Similarity	50.88%; Pred. No. 2.7e-40;
Matches		1213; Conservative	0; Mismatches 1098; Indels 78; Gaps 7;
QY			17 TTGAGGCCACTGTGAGAGACTCACCTGCAGGAGCTCACCTGCAGGAGCTGAGTCGTCGGCTGGTAC 32303
Db			TGAGGAAACTGTCGACCGSGATGAGCCTCGAGGAGCAGGNGAGCCTGTTATCGGGCAGG 32303

QY	77	ACTTCCTGGCACACAAAGGCTCTCCCCAAGCATGGAGTCCCTCTCCCTTCAGATG	136	QY	1154	CCATCTTGGGAGGAGTGGCTCACGCCAGGGCCTCCGGCATCGCTGGAGGTT	1213
Db	32304	ATTTCCTGGCTGCGGCCATCGGGTCGGGCACTCCGGTACCGAG	32363	Db	33420	CGGGTGGCACCAACCCGCTTCAGCGCTTCAGCGAGACGGCTTCACCGT	33419
QY	137	GCCCCAACGCGTAAGAGGACCAAGTCTCAGTCATGGC	193	QY	1214	TCAACGAGCCCCCTGGTACCCCTAACGCCAGCACATGAGAGCTTCCTC	1273
Db	32364	GCCCCAACGCGGCCCTGGGCGCTCGGCGGGCTCGTCAAGGGCGCTTC	32423	QY	1274	CGGACATGCCACCTGGTGCATCTACCAACCCAAAGGGCGAGACAGTGGGAC	1333
QY	194	CTTGCGCACCTCGCTGGTCACATCACAACTCTGCGAAGAGGCGT	253	Db	33480	CGTCTGGATCGGGCGACTCGGGCCAGGTGATCCACTGCATTC	33539
Db	32424	CGTGTGGGATGCCGCGTGGCGCCACTCTGGAGGTCACCG	32483	QY	1334	TGGAGGGPACGTCACGCCGAGGAGCATGCGCTCGGCTCTCGTG	1393
QY	254	TGATGGCAAAAGGGCATGGCTTAAGTGCATGTCATGGCCGACTATCAAC	313	Db	33540	TCACCGGGGTTACCCCGGAGCGGCCGACATCGCTCGGATCTACGGCC	33599
Db	32484	CTTGCGCGAGGGGAAATCGAAGGGGCCATGTGACCGTGAAC	32543	QY	1394	GCACGCAAAGGTACCTAGACGACATGGGACGAGCAGCCACAGCAG	1453
QY	314	TGCAAGGCTCCCTCTGGTCACGAAAGGGCAATTTCGAGTGT	373	Db	33600	GTTTTCCGGGTCGTCAGCGAGGACATGCTGTGCGATCGTGTG	33659
Db	32544	TCCATGTTCTGGTCACGAAAGGGCAATTTCGAGTGT	32603	QY	1454	CGCGGACCTTGGGAGGCTCGGAGGAGGACGGGAGGAGGAGGAC	1513
QY	374	CGGGCTTGGGAGCTGGGCTCTATCGGGCATCTAGACGACATGG	433	Db	33660	AAGCCGACCTCTTGGAGGAAGGCTGGCAGAGGAGGAGGAGG	33719
Db	32604	CCGCCAGCCCTGGGGTGGCTATATTCGAGGGTCAAGGAGGCG	32663	QY	1514	AGGGCAACGCTCAAGTCAGTCAGTGGCTCGACCCACCTACACCC	1573
QY	434	TCAAGGACTTTGTCATGATCAGGAGGACGGCATGAGTCAGTC	493	Db	33720	CGGGCGCGCTACGAGTGGTATCGAGTTGCGCACCAAGCGCT	33779
Db	32664	TCAAGCCTTGGCGCAATGATCGAGATCGAGTCAGCAGAC	32723	QY	1574	GGCACACATCGUCCCCGGCACGGCCTCTCGCGPRGGGTC	1633
QY	494	CGGAGGGGCTCTCCGCGAATGAACTCTACCGACTCCCTTCC	553	Db	33780	TGCGGGCTTTCGCGCCGCGATCGGCCCTGGCTGGCGAC	33820
Db	32724	ACGAGGGGGCTGCTCGTCACTGACGCGTCATGCGTGTG	32783	QY	1634	ACAGGCCAATGAAAGTCGTCGAGGACGACCCATCGTGTG	1693
QY	554	AGCCGGGTGCTCCGCGAATGAACTGCGTCAATGCGTGTG	613	Db	33821	-----CGCGGATCGCCAGGGTGGCGCTGACCCGGAGGTGG	33875
Db	32784	GCACCTGGGGATCATGAGCTCTACACAGCTACATGCGAC	32843	QY	1694	GGAGGGCTAACGCCACTGGAGGACGAGGGAGGAGGAGGAG	1753
QY	614	CTAAATATCTGATGGATCTGAAAGGAATGGGTTGGATGGCCTAATCATGAGC	673	Db	33875	TGCGGGGCCAACGGGGTGGACAGCGACGACTCGTGA	33935
Db	32844	CCTGGCGCTGCGACCGAGGTGCTGCGCAGGGATGGCACTATGACGGGTGGATG	32903	QY	1754	CCGACACATCGUCCCCGGCACGGCCTCTCGCGPRGGGTC	1813
QY	674	ACTGGTACGGCATACAGACACAGAACGCGTGTGGAGGCGCTGAGATC	733	Db	33936	CGGCGGGAGGAGGCTATGCCGGGTTGGGGTGTGACCCGGAC	33995
Db	32904	ACTGGTACGGCTGCGACTCAGCCGGCGACGTCATGCGAC	32963	QY	1814	TCATGCGAGGGGACCCCGAGGAGTGGCTCGACGCCCGCGTC	1873
QY	734	CGGACCTCCAGCTTCGAGGAGAACACCTCAAGTCACGTC	793	Db	33996	TGCGCGACCGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	34055
Db	32964	CGGGGGCCGGCCGACCGGGGGGGGGGGGGGGGGGGGGGGGG	32023	QY	1874	AGGGCTGTACGGGGCACGAGGGGAACTCCATGGCCACGGT	1933
QY	794	TTATCCTGGCTATGACGAGGGCTGGGAGTTCTCATGTCAGTC	853	Db	34055	AGGCTGGPATCGGGGAGGAGGGGGGGGGGGGGGGGGGG	34115
Db	33024	GCGCGAACGCTGCGACCCGGGTCTCGAGGTGCTGCGGCTG	33083	QY	1934	ACAAACCCCTGGGCAAGCGTCATGCGAGTGGCTCCACCGCC	1988
QY	854	CCTCCGGAGTGAAGGAGAACGGGGAGGACGACTGTCACAC	913	Db	34116	CGGACGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGGG	34175
Db	33084	CGCTGAGGACGATCACAGGAGTGGTGGAGGCGCGACGTC	33140	QY	1989	-----CGGTCTCTCACTTCGACGGAGGAGGGGGGGGGGG	2041
QY	914	CTCTCCCTGGAGGTTGGCAAGGAGCATGCTGCTGAGAGAAC	973	Db	34176	ACTGGCAGGATCGCGAACATATCCGGCTCTGACGGTCA	34235
Db	33201	TGCTCATGAGGGTGGCGCAAGATCGGGTATGGGGCACT	33260	QY	2042	TCTAGCTGGTACGGACTAGATGCGCTGGCTGAGGACT	2101
QY	1034	ACACGGGGGGCTGGGGCACTCGGGCTTACACGAGTC	1093	Db	34235	TGTCATGCGCTACGGACTAGATGCGCTGGCTGAGGAGG	34295
Db	33261	TCTATGGGGGGGCACTACCGCACTTCGCGTGGGGATGG	33320	QY	2102	ACGGCTGTCCTACACACTTGGCTTCACATCTGCGCTAC	2149
QY	1094	TGAGCAAGCAGCTGAGACAGCGCCATCGAACACCGCTC	1153	Db	34295	TGGACTTGCTACACACTTGGCTGGCTGAGGACT	34355
Db	33321	T-----GGGGGGCTGGGAGAAGGCGCTCAGCATGCG	33359	QY	2150	ACAGGAGGCGCAAGCTGAGGAGTCCCTGGTGAAGAC	2209
QY				Db	34356	AAGGGATGCGCAGTGGTGAACGGTGGTACGGTCA	34415
QY				QY	2210	CACGGGCCCCACCTCTAGTCACGCCCTCAAGGGCCAGAT	2269

Db	34416	CGGAACTGGGCCAGGCTCAAGTGTGGCGCGGATGCG--ACGGTGTGGCGCGGCCA	34472	Db	8857	GATCTACTTATCAGGCTTCAGCITAGTGTATCTGAGATGCAAGTAAC	8916
Qy	2270	AGAACCTCAAGGCCCTCGAAAGGCTCAACTGAGGCCGGAGAAGGGGTGACAA	2329	Qy	421	GTCCAGCTTACCAATCAAGCCACTTTCGCAATGTCAGGAGACAGGCCATGATGGTG	480
Db	34473	AGGACCTGAGCTTCTGAGAACGCACTGAGCCGGGGAGGCGAGCGCGTGACCC	34532	Db	8917	GTCCAAGCATGCAATTAAAGCTTGTGTTATGATGATGAGATGAAAGGAACTCTGTC	8976
Qy	2330	TCGAGGAGCAGAGAGTACGTCTGCTGCTATTGATAGGAGCGGGAA	2378	Qy	481	CAGAGCATCGTCAGGGGGCTCCCGTGAATCTACGGACTCCGTCCAGATGCT	540
Db	34533	TGGAGCTCGATGACGGGCTTCCTACTACCCACCGAAGGCCAGCA	34581	Db	8977	AGTATTGAGCTTCACAAAGGCACITAGAGAAGTATCTATGCACTTCAGTGGCA	9036
RESULT	3			Qy	541	GTGGAGAGCTCCAGCGGGTGTATGAGGGTACATGGCATTCAUTGGGTTGTCG	600
SPU3009				Db	9037	TGCAAAATTCATAATTCAACTCACTATGACTCTACACAGGTGATGGTACAC	9096
LOCUS	SPU3009			Qy	601	TGCAGCAGAACCTAAATCTGTGATGGGAGCTCGAAGGAATGGGTTGGATGCC	660
DEFINITION	Schizosaccharomyces pombe	39305 bp	DNA	Db	9097	GTTCCCAATCTCGAAATTATTAGACAAATTAAAGAGAAAGTGGACTGGCA	9156
ACCESSION	U3009	39305	linear	Qy	661	CTAAATCATGAGCAGTCGGCACATACAGTACCCAGAACGGCTGTGGCAGGCCTC	720
VERSION	U3009.1	GI:106316	PLN 14-Nov-1995	Db	9157	ACTATAATTCATGATGGTTGACTGTTGAGCTATCTGTTAGAAGGCT	777
KEYWORDS	Schizosaccharomyces pombe.			Qy	721	GACCTCGAGATGCCGAGAACGGAGTTGAAATGTCACACACACACTG	777
ORGANISM	Schizosaccharomyces pombe			Db	9217	GATTGGAGATGCCAGAAAGCGAGTTGAAATGTCACACACACTG	777
REFERENCE	1 (bases 1 to 39305)			Qy	778	TCCAAACSGAACCCCTTATTCACGCGTATCCACGAGCTTCGGAGGAG--AACACTCAACTTCACACAC	837
AUTHORS	McCombie, W.R.			Db	9277	GGATCCAAGAGCTTGGAGTCATTGAGTAAATTTAGAAGAGAAAGTGGCT	9336
TITLE	Sequence analysis of a region of the fission yeast genome			Qy	838	GTCAAAGAAGTGTGCTCCGGAGTGACGGAGACGCCCGAGACGACTGTCACACAC	897
JOURNAL	Submitted (02-AUG-1995) W. Richard McCombie, Cold Spring Harbor			Db	9337	GTAACACATTCTTGGAGTCATTGAGTAAATTTAGAAGAGAAAGTGGCT	9396
FEATURES	Unpublished			Qy	898	ACCCCAAACGGCAGCTCCCTGGCGAGAGGTGATGAGGAGGCTGCTGAAAG	957
SOURCE	2 (bases 1 to 39305)			Db	9397	TCTTGTGTTAGAGAAGC--ACTTAAATTTGCAACCGTCATGTCAGTGAAG	9453
AUTHORS	McCombie, W.R.			Qy	958	AACGAGAACAGCTCTGCCCTGACAAAGAGCTCCTCCAGGAGGACAGCTGCACT	1017
TITLE	Direct submission			Db	9454	AATAAAAGAAACTTTGCACTGTGAAAGGAGGGCTGCACTTATGGACCTAAC	9513
JOURNAL	Submitted (02-AUG-1995) W. Richard McCombie, Cold Spring Harbor			Qy	1018	GCCAAGCAGGCCACATACCAGGGCGAGGCTCTGGCCACTCAGGCCTACTACCGAC	1077
FEATURES	Unpublished			Db	9514	GCAAAAGTTGTAATTACAGCGGAGGTTGAGCGCAATTGAACTCTTATAGTCG	9573
SOURCE	1 (bases 1 to 39305)			Qy	1078	ACTCCCTTGAGGCCCTCAGAACGAGCAGCTCAGAGCCGGCATCGTACACGTCGGGCC	1137
AUTHORS	McCombie, W.R.			Db	9574	AGTATGTTGATGAGTCAGCAGCGAAATAGCGGCTCCAGATPATGTCCTGTTG	9633
TITLE	Sequence analysis of a region of the fission yeast genome			Qy	1138	TACACCT----ACCGTTCGCCCATCTAGCGACAGCTCCACGCCGACGGCT	1191
JOURNAL	Submitted (02-AUG-1995) W. Richard McCombie, Cold Spring Harbor			Db	9634	CATACTATCTAACCTTCAATATTCGAAATTATTAATCAACCCATGAAACCGCAAG	9693
FEATURES	Unpublished			Qy	1192	CGGGCATGCGCTGGGGCTCAAGGACGCCCTGGTACCCCTAACCCACACAT	1251
SOURCE	2 (bases 1 to 39305)			Db	9694	CATGGATACAGTGTGCAAGTTATCTGATGACCTGCACTCCGAAATAGACTTTGTA	9753
AUTHORS	McCombie, W.R.			Qy	1252	GACGAGCTCTTCACCAAGACGGCATGACCTGGGACTCTACCCACCCCA--G	1308
TITLE	Sequence analysis of a region of the fission yeast genome			Db	9754	GACGATTTAGAGTGTGAGGAGTGTGTTTACCAATTATGTAACGAAAC	9813
JOURNAL	Submitted (02-AUG-1995) W. Richard McCombie, Cold Spring Harbor			Qy	1309	GGGGCACACCTGGTACCGCAGATCGAGGAGCTACCCGCGACGAGGACTGCA	1368
FEATURES	Unpublished			Db	9814	AAGGATGGCTTCAATATAGACATGAGGTTTATTAATCCGAGGACCGCTG	9873
SOURCE	2 (bases 1 to 39305)			Qy	1369	TACGACCTGGCTCCGCGACGCCAGGCTACAGGAGCTGAGGACGCCAGCTG	1428
AUTHORS	McCombie, W.R.			Db	9874	TATGAATTGGATATCTGTGTTGGAAACGCCCTCGCTGTTATGATGATGTC	9933
TITLE	Sequence analysis of a region of the fission yeast genome			Qy	1429	GTGACACGACCAACGGAGCTCCGGGATGCTCTGCGCACCGCGAG	1488
JOURNAL	Submitted (02-AUG-1995) W. Richard McCombie, Cold Spring Harbor			Db	9934	ATCGACACAAAGACAACCCCTACGAACTATTTGAAATTGAAACATAGAA	9993

Qy	1489	GAGACGGCGCATAATCTGCAAGGCAACAGTACAGTCAGATCGAGTCGGC 1548	htp4; glucose transporter protein; glycosyl hydrolase family 3;
Db	9994	GAGAAATCTATTTTAAAGAGGGAGAAAGTATGTACGGGTGAGAT 10053	inosine-uridine preferring nucleoside hydrolase; isocitrate lyase (EC 4.1.3.1); MFS allantate permease; MFS inorganic phosphate transporter; MFS transporter of unknown specificity; NC1 allantate transporter; TEF-type LTR; transcriptional activator; zinc finger.
Qy	1549	TCCGACCCACTACACCCCAAAGGGACACATGTCGGGACAGTGTCCCGCACGCTCTCCC 1608	fission yeast.
Db	10054	AGTGACGCCACTACACGGTCAACTAATCTGTCGGGAGAGATACTCT 10113	transporter; TEF-type LTR; transcriptional activator; zinc finger.
Qy	1609	GTCGGGGCTGCAAGGTATGACGACGGCCATGACGACGGCGAATCGAAAGTCCGTGCGCTCGCC 1668	transporter; TEF-type LTR; transcriptional activator; zinc finger.
Db	10114	ATGGGTGTTGTTAAGTATGTTATGATCTGAGATGACTGAGTGTGAGGGTGGTGC 16173	fusion yeast.
Qy	1669	AGGACACGACGACAGGCTCATCATCTCGCGGGCCTAACGCCACTGGGAGACGGAGGG 1728	htp4; glucose transporter protein; glycosyl hydrolase family 3;
Db	10174	ATACAGTTGTTGTTGTTGATTTATGTTGTTGACAGGAGATGACTGAGGGTGGTGC 17233	inosine-uridine preferring nucleoside hydrolase; isocitrate lyase (EC 4.1.3.1); MFS allantate permease; MFS inorganic phosphate transporter; MFS transporter of unknown specificity; NC1 allantate transporter; TEF-type LTR; transcriptional activator; zinc finger.
Qy	1729	GCGGACCGGAGCATGAGCTCCGGGGTGCCTGACAGGAGATGACTGAGGGTGGCC 1788	fusion yeast.
Db	10234	AGAGATGAAACAAACAGTACTTACCATCGTTATCGACAACTGTTATTCTATCTG 10293	transporter; TEF-type LTR; transcriptional activator; zinc finger.
Qy	1789	GCGGACCGGACCCAAACCGTGGCTCATGAGACGGGCCACCCGGAGGAGATGCC 1848	fusion yeast.
Db	10294	CAATCCATCCATCGTCACTAGTGTGACTCATCGGAGTGGCTTATGGCTCG 10353	htp4; glucose transporter protein; glycosyl hydrolase family 3;
Qy	1849	CTGACGCCACGCCCGCTGATCCAGGCGTGTAGGGCAGACGGGAACCC 1908	transporter; TEF-type LTR; transcriptional activator; zinc finger.
Db	10354	ATCTCTGAAGCCACCTTATGCACTTGTGACATTTGTTGACACGGTAACTG 10413	fusion yeast.
Qy	1909	ATGCGGACGCTGCTCTGCGACTACAAACCCCTCGGGCAAGCTCTCCCTAGGT 1968	htp4; glucose transporter protein; glycosyl hydrolase family 3;
Db	10414	TTAGCAACATTATTGAGAGAACCCCTGTTCACTTCCGACCGAGGCCGGACGCTG 2028	transporter; TEF-type LTR; transcriptional activator; zinc finger.
Qy	1969	AGCGCTGAGACACCCGGTCTCACTTCCGACCGAGGCCGGACGCTG 20473	fusion yeast.
Db	10474	AAAAGTGTGAGATATCCGCTTACATTGTTGTCAGCGGTGCGCATGTT 10533	htp4; glucose transporter protein; glycosyl hydrolase family 3;
Qy	2029	TACGGCAGGAGCTACTGCGGTACAGGACTACAGGTGTCGACAGGGCTAAT 2088	fusion yeast.
Db	10534	TATGGAGAAGAGCTCTGTTGGTATAAGTATTATGAAAGCCGTAGAAAGGG 10593	htp4; glucose transporter protein; glycosyl hydrolase family 3;
Qy	2089	TTCCCTTTGGCACGCCCTGTCCTACACCACTTGTGCTTTCAATCTCCGTCT 2148	fusion yeast.
Db	10594	TTTCCATTGTTGATGGCTTCTATACATCTACTTGTGAAATTCCATCTTATTGAG 10653	htp4; glucose transporter protein; glycosyl hydrolase family 3;
Qy	2149	CACAAAGACGCGAACGCTGAGGGCTCTCGCTGAGAACACCGGTCTCGTGC 2208	fusion yeast.
Db	10654	AAATGGTGTGAAAGTTAAGGTGCAATTGCGTACTGGCCCATGCTGT 10713	htp4; glucose transporter protein; glycosyl hydrolase family 3;
Qy	2209	GCACAGGGCCAGCTACTGCAAGGCCCTCAAGGCCAAAGATAACGCCGCC 2268	fusion yeast.
Db	10714	GGAGAAATTATACAAGTATATTTCTCAATC --GTTCTGTTCAATCAGCTGT 10770	htp4; glucose transporter protein; glycosyl hydrolase family 3;
Qy	2269	AGGGAGTCACGGCTCGCAAGGTGGAAGTCAGTCAGCCGGAGACGAAGGGTGAC 2328	fusion yeast.
Db	10771	AAAGAACTAAAGAGTTCTAAAGACCAATTATTTAGA 10830	htp4; glucose transporter protein; glycosyl hydrolase family 3;
Qy	2329	ATCGAGGAGCAGAGAAGTACTCTCGCTGCTGTTGATGAGGAGCGGGATCTGGGT 2388	fusion yeast.
Db	10831	ATGGAGTTGATATAAGTATCTCTTATGATGACTTATGAAATGGGT 10890	htp4; glucose transporter protein; glycosyl hydrolase family 3;
Qy	2389	GTCGAAAGGGGACTATGAGGTATCTGAGGAGCAGCGCAG 2434	fusion yeast.
Db	10891	TCTGAGAAGGGGAGTACAACCTCTCTGCGAACAGCAAG 10936	htp4; glucose transporter protein; glycosyl hydrolase family 3;
KEYWORDS		adenosine deaminase; alpha-glucosidase; beta-glucosidase precursor	
RESULT 4			
SPBC1683	SPBC1683	41799 bp DNA linear PLN 23-AUG-2000	
LOCUS			
DEFINITION			
ACCESSION			
VERSION			



Oy	301	CCGACTATCAACATGCAACGCCCTCCCTCTGGGACGCGCTCGAGTCGATGGTGA	360	Oy	1369	TAGGAGCTGGCCCTCGTCGTCGCGGACGSSAASCGTACGACGACAGCTGTC	1428
Db	11635	CCAACCGTTAAATATCCATCGGACCTTAAATGGTGTGGGTTTGAATCATTTCTGA	11694	Db	12712	TATGAATTGGATATCTGTTGGACGGCTCTGCTGTGTTATTGATGATGTCATTA	12771
Oy	361	GATCCGTCTGGCGGGCTGGAGGTGCGCTCATCGGGGATTAGAGACTGGA	420	Oy	14249	GTGCACAGGCCAACGCCAGGCTCCGGGATGCCCTCTGGCTCCGCCAUUCGGAG	1488
Db	11695	GATTCTTATCAGTCAGTCAAGCTTGTGGAATATGATCAGAGACAGGGCATATGGG	480	Db	12772	ATCGACACAAAGACAACGAAACCCCTACGAACTCATCTTGAATTTGAGACAATGAA	12831
Oy	421	GTCGAGGCTACGATCAAGCTTGTGGAATATGATGATGGCATATGGTGTG	480	Oy	1489	GAGACGGCGCATCACTCGTCGAAGGGAAACACGTCACAGTCAGATGGTGGC	1548
Db	11755	GTCCAAAGCAAGTAAAGCATTTGTTGTTGATGATGGAGATGAAGGACTCTGTC	11814	Db	12832	GAGGAAATCTTATTAAAGAAGGGGAGAAGTATAGTACGGTCAAGTGGT	12891
Oy	481	CAGAGCATCGTCACGGCGGCCCTCCGGAAATCTACGCCATCCGGTCCGAGATTGCT	540	Oy	1549	TCCGACCCACCTACACCTCAAGGGCGACACATCGTCCCGGCCACGGCTCCCGCC	1608
Db	11815	AGTATAGACGCTTCACAAAGGGCAGTTAGAGAACTATCTTGTGGCATTTGGCA	11874	Db	12892	AGTCACGGCACTTACACCGTCACTAATCTGTCGCCAGTACGGAGHAGATCTC	12951
Oy	541	GTGCGAGACTCCAGCGGGGTGGCTCATGAGGGTACATGGCATCATGGGTG	600	Oy	1609	GTGGCGCTGCAAGGTCATTCAGCAGCAGGGCGAATCGAAAGTGGATGACAC	11934
Db	11875	TGCAAAATATTCAAATTCAAAATCAGTCACTTACAGTCTTACACAGGATGGTGA	11934	Oy	1669	AAGGAGCAGGAGCTCATCTGGCCCTTACGCCAGCTGGAGACCGAGGGC	1728
Oy	601	TGCAAGGAGAACCTTAATAATCTTGTGGATGCTGCAAAAGGAATGGGTGGATG	650	Db	13012	AAATCAGTGTGTTGATGTTGTTGTTGCGGTTGAGCAAGAATGGAAACG-SAGGT	13071
Db	11935	GTTCCTCAATCTGAAATTATAGACAAATTAAAGAAGAGTGGAGTGGAAAGGT	11994	Oy	1729	GCGGACCCGGCGACATGAGCIPCGGGGCGCTGACGACCGACCTCAGTCGAGG	1788
Oy	661	CTAATCATGAGCGACGNGTACAGTACAGTACAGTACAGTACAGTACAGTACAGT	720	Db	13072	GAAGATAGAAACAACTAGCTTACATCTGTTATCAGCAACACTGTTCTG	13131
Db	11995	ACTTAAATATCTGATGTTGTTGAACTTATCTGTTGAGAACTATGAGCTGAGGT	12054	Oy	1789	GCGGGAACCCAAACACCGCCATCTGGCCATGAGCAGGGCTAGCCAGTCTGAG	1848
Oy	721	GAACCTGAGATGCCCGACCTCACCTCCGGAGGAG--AACACTCTGTCACACT	777	Db	13132	CAATCCAAATCCAAATACAGTACAGTACAGTACAGTACAGTACAGTACAGT	13191
Db	12055	GATTTGGAGATGCCAGGAAGCCGAGATGCAAGAATGTCACACTATCACTGTA	12114	Oy	1849	CTCGACGCCAACACCGCCATCCAGCCTGCGCATCGACAGACGGCACCCAGGAG	1908
Oy	778	TCCAACGGAAAGCCCTATCCACGCTATGACAGAGGGCTAGGGAGTCTTCAGTC	837	Db	13192	ATCCTGTGAGCCACACTATGACATGACATTGACATGACATTGACAGTAAAGT	13251
Db	12115	GGATCCAAGAGCTTGGAGCTTACATTAGTAAGAAGAACAAAGCAAGAATCTGCA	12174	Oy	1909	ATTCGGAGCTCTTGGGACTACACCCCTCGGCAAGCAGCTGCGCTCAGTGTCCC	1968
Oy	838	GTCAAGAAGTGTGCTCGGAGTGAACGCGCCCGAGACGACTGTCAAACAC	897	Db	13252	TTACAAACATTTGAGGACCAAACTCTGCGCATACATTCATCTGCGCATAC	13311
Db	12175	GTGAAACATCCTGGAGAAATACAGAACAGAAATCATGTCAGTAAATGATCC	12234	Oy	1989	AACGCCCTGAGACACCCCGCTGCGCATCGACGGCACCCAGGAGCTGCGCAACTC	2028
Oy	898	ACCCCGAACGGCACTCTCCGGAGGTGGCAACGGAGGACATGTCGCTGAG	957	Db	13312	AAAACACTGAGGATAATCTGCTCATCTTCATCTGTCAGCGGTGCGCATGTT	13371
Db	12235	TCTGTTTAAGAAGC--ACTTAAATTCGAAAGGAAAGAGACGCTGAGCTGAA	12291	Oy	2029	TACCGCGAGGACCTCTGCGGTACAGGAGCTACCTACCTCGCACGGCGCGCGCAG	2088
Oy	958	ACGAGAACAGAGCTGCGCTTGGCAAGAAGAAGAGACGCTGAGTGTGCGCCAC	1017	Db	13372	TATGGAGGAGGACCTTGTGGTATGAGTATTAGAAGCGTAAAGGAGGCTTG	13431
Db	12292	ATAAAAGAACACTTGTGCAATGCAAGGAGGGCACGTTGAGTATGACCTAC	12351	Oy	2089	TTGCCCTGGGACGCGCTGTCGCTACACCACTTGGCTTCCCTTCACATCTGCT	2148
Oy	1018	GCCANGCAGCCACATACACGGGGGGCTCTCCGCACTCAGGGCTACTACGCACTC	1077	Db	13432	TTTCATTTGATGGCTTCTACATACACTTTGATGATTTGATGATTCATCTT	13491
Db	12352	GCAGAAGTTGTTAATACAGCGGAGGTGGCTGGCAATTGTAAGCTTATTACTG	12411	Oy	2149	CACAAAGGGGAAAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG	2208
Oy	1078	ACTCCCTTGACGGCTCAGCAAGGAGCTGAGACGCCCAAGTGTACACGGGCC	1137	Db	13552	GCAGAAATTATACAGAATTTCTCAATC---GTTCTCTGAGTAACTGACCTG	13608
Db	12412	AGTATGATGATGCACTACAGGAAATAGACGGTGTGCAAGTGTCTGGTGC	12471	Oy	2269	AAGGAGCTGAGGCTGAGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG	2328
Oy	1192	CCGGCGCATGGCTGGGGCTCTCACAGGCCCTGGTACCGCCACATT	1251	Db	13609	AAAGACTAAAGGTTTCTGAGCTTGTGCTTAAAGGACCAATTATTAGA	13668
Db	12532	CATGGATACGTTGCGAAGTTATCTGAACTCTCCGAAATAGACTTGTATA	12591	Oy	2329	ATCGAGGAGGAGGAGGAGCTGAGCTGAGCTGAGCTGAGCTGAGCTGAG	2388
Oy	1252	GACGAGCTCTTCACCAACGAGACATCCACCTGGAGACTACACCCCAA--G	1308	Db	13669	ATTCGAGTGTGATGATGATGACTCTTGTGAGTGTGTTATGATGAAATGGT	13728
Db	12592	GACGTTACGTTAGAAGATGAGTGTGTTTGTGATGATGACGAAATG	12651	Oy	2389	GTGAAAGGGTCACTATGAGTTCTGAGCGACGCCGCGCAG	2434
Oy	1309	GCGCAGACCTGGTACGCCGACATGGAGGAGCTACCCGCCGAGGAGCTGAC	1368	Db	13729	TCTGAAAGGGGAGTACAGTGTCTGCGAACCGAGCAGTAAG	13774

RESULT 5								
SP33010	SPU33010	85837 bp	DNA	linear	PLN 14-NOV-1995			
LOCUS	Schizosaccharomyces pombe	cosmids	359, 1198 and 1683.					
DEFINITION	Schizosaccharomyces pombe							
ACCESSION	U33010							
VERSION	U33010.1	GI:1663617						
KEYWORDS	Schizosaccharomyces pombe.							
SOURCE	Schizosaccharomyces pombe.							
ORGANISM	Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes; Schizosaccharomycetidae; Schizosaccharomyces pombe.							
REFERENCE	McCombie, W.R.							
AUTHORS	McCombie, W.R.							
TITLE	Sequence analysis of a region of the fission yeast genome							
JOURNAL	Unpublished							
REFERENCE	2 (bases 1 to 85837)							
AUTHORS	McCombie, W.R.							
TITLE	Direct Submission							
JOURNAL	Submitted (02-AUG-1995) W. Richard McCombie, Cold Spring Harbor Lab, 1 Bungton Rd, P.O. Box 100, cold Spring Harbor, NY 11754, USA							
FEATURES	source							
BASE COUNT	27560	a 14781 c 15395 g 28101 t						
ORIGIN								
Query Match	15.7%	Score 393.6; DB 8; Length 85837;						
Best Local Similarity	49.0%	Pred. No. 7.2e-39;						
Matches	1199;	Conservative 0; Mismatches 1229;	Indels	18;	Gaps 5;			
QY	1	ATGCTGTATTATGTTGAGTTGATCAATTCAATGGATTAAGTGGAGGCTATG 60						
Db	29116	ATGATGGAACATGATGTTGAGTTGATCAATTCAATGGATTAAGTGGAGGCTATG 29175						
QY	61	CTCTCTGGCGTGTGACTTCGCACTGCTGCAACAAAGCTCTCCCAAGCATGGCTTGCGATCTGCTCA 120						
Db	29176	CTGCTTCTCGGCCTGACTGACTCTGCTCACTGCAAACTTCCGTCTTCA 29235						
QY	121	CTCCGCTTACAGCATGGCCACAGGGCTAAGGGGACAACTCTCAATGGGRCCT 180						
Db	29236	ATTAGATATCACAGCGCCCTAAATGGTATCGTGGAACTAGTAGCTCTCAATTCTCACCA 29295						
QY	181	GCGGCTCTGCTCCCTGCGCACTGCTGCGTGGTCCACATTACCAACACTCTGCGAA 240						
Db	29296	TCCGCTTGTCTTCTCTGCGGACAGCCGTAGGGCTACTTCTGCGATGTGCTCCGGC 300						
QY	241	GAGGCAAGGAAAGGATGGCAAGAGGCCATGCTCTAGAGTGGCGCATGTGCTCGAA 240						
Db	29356	GAAGTGGGAATTTAACAGAGAAGAACAAAGCGAAGGGTGTAGTGTGTTTGGT 29415						
QY	301	CCGACTATCAACTGCAACGCTCCCTCTGGGCGACTCTGAGTCGATGGAG 360						
Db	29416	CCACCGGTTAAATTCATCGTGGACCCCTTAATGGTAGGGTTGTGATCATTTCTGAA 29475						
QY	361	GATCGTTCTGGGGCTGGGACTCTCATCGGCCATCAGGCACTGG 420						
Db	29476	GATTCTACTTATCAGGCTTGGAGCTAGTTGATCTGATGCAAGTAAAC 29535						
QY	421	GTGCAAGGCTTACGATCAAGGACTTGTGCAATGAGTCAAGGAGACAGGGCATGATGTG 480						
Db	29536	GTCCAGCATGCAATTAGTTGATGATGATGAGTCAATGAGAAGGACTCTGTC 29595						
QY	481	CAGACATGCTCAGGCGGGCTCCGTGAATCTGCACTCCGGTTCCGATGCT 540						
Db	29596	AGTATGACCTCTACAAAGGGACTTAGAGAAGTATCTTATGCAATTCTGTC 29655						
QY	541	GTGCCAGATCCAGCCGGTGTCTCATGACCGCGTACAATGCACTAATGCCCTCG 600						
Db	29656	TGCAATATTCAATTCAACTCACTTAACTTACGACTCTCTGACAGGATGGACAC 29715						
QY	601	TGCACGGAGAACCTTAATATCTGATGCGATGCGAAAGGATGGGGTTGGATGCG 660						
Db	29716	GTTCCCAATCTGAAATTAGACAATATTAAAGAAGAGTGGGGATGGAAAG 29775						
QY	661	CTATCATGAGGAGCTGGTACGGCACATACAGTACACAGTACACAGAAGCGCTGCGC 720						
Db	29776	ACTATATCTGATGTTGAGTGGACTTATCGTTGAAGAAGCTATGAGTGCAGT 29835						
QY	721	GACCTGGAGATGCCGGACCTCAGCTCAGGAGGAG -- AACACTGAGTCACGTC 777						
Db	29836	GATTGGAGATGCCAGGAAGCCGAGATTCGAATGTCACACTATCAACACTATG 29895						
QY	778	TCCAACGGAAAGCCCTTATCCACCTCTGCGAGAATCTGAGGAGCTGAGGCT 837						
Db	29956	TGCAACATTCCTGCGAGAATCTGAGGAGCTGAGTGGAGTGGAGCT 29955						
QY	898	ACCCCGAAGAAGCAGCTCTCTGGGAGTGGCAACCGGGCATCTGCTGCTGAG 957						
Db	30016	TCTTGTGTTAGAGAAGC---ACTTAAAAATTGCAAGCCAGTCATGTCCTG 30072						
QY	958	AACGAGAACACAGCTCTGCGCTTGAGCAACAGAAGAGGAGGAGGCTGAG 1017						
Db	30073	ATAAARAAAGAAACTTGTGCAAGGAAGGCCAGTGGAGTATTGGACCTAC 30132						
QY	1018	GCCAAACGAGGCCACATACCAGGGCGAGGCTGCGACTCAGGCCACTAGGCC 1077						
Db	30133	GCAAACTACATTCCTGCGAGAATAGCAGAAATCATGTTGAGTAAATGATTC 30192						
QY	1078	ACTCCCTTGTGCGCTCACCAAGGAGCTGAGTGGAGGAGGCCATGTACAC 1137						
Db	30193	AGTATGTTATGATGGCATAGCAGCAAAATAGACGGGTGTCGAGTATGCTG 30252						
QY	1138	TACACC---ACCGTCTCCCATCTGAGGAGGTGGTGGAGCGGC 1191						
Db	30253	CATACTATCTAACCTAACATATTGCAAAATTATTCACACCTAGAACCGGA 30312						
QY	1192	CGGGCATGCGCTGGGGCTCAACGACGCCCGTGGTACCCCATACGCC 1251						
Db	30313	CATGGATACCTGTTGCGGAAGTTTATCTGAGCTGACCTCCGAAATAGACCTTGATA 30372						
QY	1252	GACGAGCTCTTCAACAAACGGACATCACCGGCTACCCGGAGCGAGGACTGCA 1308						
Db	30373	GACGATTAGCATTTAGAGTGGAGTGTGTTACGATTATGTAAGACAAATG 30432						
QY	1309	GGGGCAGACCGTGGPACCGGACATGGAGGGCACCTACCCGGAGCGAGGACTGCA 1368						
Db	30433	AAGGATGGCATTTCTATAGACGARTGAAGTTATTAACTCTGACGAGGACCTG 30492						
QY	1369	TACGAGCTGGGCTGTTGCGGACGCGACGCGAACGGAGCTAGGAGCAGCAGCTG 1428						
Db	30493	TATGAATTGGGATACCTGTTGGACGCTCTGTGTTATGATGATGTCATTAA 30552						
QY	1429	GTGCAACAGCCACCAAGCAGTGGCCATGCGCTCTGTTGCGACCCGGAG 1488						
Db	30553	ATGACACACAGACAAACCCCTACGACATTAATTGTAAGCTGACATG 30612						
QY	1549	TCGGCACCCATACCCCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1508						
Db	30673	GAGACGGCGCATCACTGTCAGGGACAGTACAGTCAGGAGGAGGAGGAGGAGG 1548						
QY	1550	GAGAAATCTCTTATTAAAGAAGGGAGGAAGTATGTTACGCGGAGGAGTGG 30672						
Db	30674	TCGGCACCCATACCCCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1508						
QY	1551	TCGGCACCCATACCCCTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1508						
Db	30733	ATGGCTGTTAAGCTTGTGTTGAGGAGTAACTTGTGAGGAGGAGGAGGAGG 30792						



QY 554 AGGGGGTCCGTTCAATGAGGCCATCAATGGCATCAATGGCTGCGAGGAGACC 613  
 Db 612 CCCCATGGGCATCATGACGCCATACACCGAGCTGAAAGGGCTCATGGAGAAC 671  
 QY 614 CTAATATCTGATGGATGCTGAAGGAATGGGTGGCTAATGGCTAAATGAGGG 673  
 Db 672 CCTTCTCGTCAGCAGGCTCTGGTAAATGAGGAATACAGGGATGAGTC 731  
 QY 674 ACTGCTACGCCACATACATACACAGAGGCCCTGGCAGGCCGACCTGGAGTC 733  
 Db 732 ACTGCTCGCCATGATCAGCGTGTATGTTCAANGCCGCTCTAGACCGAATG 791  
 QY 734 CCGAACCTCAGCAGGCTCCGGGA--GAAACACTCAAGTCAACGTCMCAACGGA 790  
 Db 792 CTGGCATATAATGGCTACACTGTGATCTAGTCACCGTACTATTCAGCTGC 851  
 QY 791 CCTTATCCAGCTATGGCAGAGGGCTGGGAAGTCTTCAGTCTCAAGAGGTG 850  
 Db 852 TCAACCCAGCAGGCTCCGGGA--GAAACACTCAAGTCAACGTCMCAACGGA 911  
 QY 851 CT-----GCCTCCGGAGTACGCGAGCCCGAGACGACTGTCACACACCCG 904  
 Db 912 CTCAGGAGGCCAGAGATCTCGCATGSGCATGGAGCTACAGCTGAGCTTG 971  
 QY 905 AACGGCAGCTCTCCGGAAAGGTGGAACCGAGGGATCGCTGTCGAGAACG 964  
 Db 972 CGCAGACGCTGCTGCGGCATCGCATCTGAGTCATCGCTCAAAATGTA 1031  
 QY 965 ACAACGTTCTGCCCT-----TGAGCAAGAAGAAGAAGACCTGATGTGG 1015  
 Db 1032 AGTTCTCTCTTAACCTGAGGAGGCCGATCAAGAGATGCCGCGTGTGGCA 1091  
 QY 1016 AGCCAGCAGGCCACATACACACGGAGGCTCTGCCCACTCAGGGCTACTAC 1075  
 Db 1092 ATGCAAGGCGCAAGTCTGTCGCCGGAGGGCTGCCGGGCTGAGGGCTACT 1151  
 QY 1076 TCACTCCTTGACGCCCTGAGCAAGCAGCTGAGACGCCGACAGTACAC 1128  
 Db 1152 TCTCACCTAGACGCTTCTGGCGCTTCTGGCGGCTCTGGCGCTACATCA 1211  
 QY 1129 -----GTGGCCCTACACCCGGCTCTCCCATCTGGCAGCTGGCTCACCC 1180  
 Db 1212 GTGAGGTTGCTGAGCTATAAACACTACCCACCTCGATGGATCTGAAACCG 1271  
 QY 1181 CGCACGGCGCTCGGGCATGCCATTAGGGCGCCCTGAACTCTGAGTCATCA 1231  
 Db 1272 CGGGCGACGTGATGGATGGACCTGGCACACCCACGACTGGCACAGCTG 1331  
 QY 1232 CCTCTAACCGCCAGCACATGGAGCTCTTCTACCAAGAGGACATCACCG 1291  
 Db 1332 CTCCTGACCGACCATCTCGTCCCGCATGGTGGACGAAACCAATCTCATCG 1391  
 QY 1292 ACTACTACACCCAGGGGAGAACCTGGTACCCGACATGGAGGG-----CAGTACA 1348  
 Db 1392 CCTCGTATCCCAAGGGATTACAAAGGGTGAGGATGAGGCTACCTGAGC 1451  
 QY 1349 CCCCCGAGGAGACTGGACCTACAGGCTCGGCCCTGCTGGCACGCCAAGGG 1408  
 Db 1452 CAGAGAGAGAACACATTGAACTGGCTATGGCTGCTGCGGCTAAGCT 1511  
 QY 1409 ACTTAGAGGACGACTGGCTGAGCACGCCACCGAGCTCCGGCATGGCTC 1468  
 Db 1512 GGGTTGAGGAGACTGGTAATGACACTGGACCGCTGAGCTGCTGGTGAAG 1571  
 QY 1469 TCGCTCGGCCACCCGGAGGAGACGGCCGCACTCATCTGTCAGGGACACG 1528  
 Db 1572 TTGCACTGAGTACTCAGGAGAACCTGGTATCTCTCAAGGTGGCAGGAGCT 1631  
 QY 1529 AGTCAGATCGAGTGGCTC-----CGCCCCACTACCCCTCAAGGGGACCA 1582  
 Db 1632 AAATCTATGCTGAGTACTGTTAATGTCGCGCCTGGCTGATGAGGCGCA 1691  
 QY 1583 TCGTCCCGGCAAGGCTCCCTGGCTGGGGCTCAAGGCTATGACGCC 1642

RESULT 7  
 AF005277 AF005277 3578 bp DNA linear BCT 02-MAR-1998  
 LOCUS Cellulomonas biazotea cellobiase (cba) gene, complete cds.  
 DEFINITION Cellulomonas biazotea  
 ACCESSION AF005277  
 VERSION AF005277.1 GI:2921739  
 KEYWORDS SOURCE  
 ORGANISM Cellulomonas biazotea  
 Bacteria; Firmicutes; Actinobacteria; Actinomycetidae;  
 Actinomycetales; Micrococcineae; Cellulomonadaceae; Cellulomonas.  
 REFERENCE 1 (bases 1 to 3578)  
 AUTHORS Wong, W.K., Ali, A., Chan, W.K., Ho, V. and Lee, N.T.



		MEDLINE	93165721
		PUBMED	1537792
	FEATURE	source	Location/Qualifiers
Db	1884 AGCGTACCGCGACGGGACGGACGGACGGCGACGGCGACGGCGCGCGCGCGCGCGCGCGCTC	1943	1. .3710
Qy	1594 CAGGGC-----TCCTCGGRRGGCGCTGCAAGGACACGACAGGTCATCATCGACGACCGCCGAA	1644	/organism="Agrobacterium tumefaciens"
Db	1944 GCGGGCGCTCGGACCTGTTGGGGCTCGTGGCGACGGGCGACGGCGCGCGCGCGCGCGCGCGCTC	2003	/cell_line="33/73"
Qy	1645 ATCGAAAGTCGTCSCCTCGCAAGGACACGACAGGTCATCATCGACGACCGCCGAA	1704	/tissue_1b="genomic HindII cosmid"
Db	2004 ATCACCGCGCCGCGCCGCGCCGCGACGGGACGGTGCCTCGTGGCGCTCGTGGCGCTCG	2053	169. .2637
Qy	1705 AACGCCACTGGAGAGGGGGCGACGGGACGGGACGGGACGGGACGGGACGGGACGGGACGGCTC	1754	/gene="cbg1"
Db	2064 ACCGAGCAGGAGAGGGAGCTCGCGACAGTCACGATCGCCCTCGCCGGCGAG	2123	169. .173
Db	2124 GACGCCTCGTGGGGCGTCGCCGGGCCCTCGCCGGCGACGGTGTGGGGCTCGAGCC	2183	/product="cbg1"
Qy	1825 GGCACCCCGAGGAGATGCCCTGCGCTGACGCCACGCCACGCCACGCCACGCCACGCCACGCC	1884	/protein_id="AAZ22082.1"
Db	2184 GCGACCGCGGTTCTCAGCGCTGCGACGGACGTCGACCGCGCTGCTGGGGCGCGTC	2243	/db_xref="GI:12222082"
Qy	1885 GCGGCACGAGACGGGAACTCCATGCGGAGCGTCTTGGGACTACAAACCCCTCG	1944	/translational="MIDDLEKMLEQEVSLISGADFWTTAERLGVPKIKVTDQPN
Db	2244 CCGGAGCTGTCCTCACTTCCAAAGCGCTGAGGACACCCCGCTACTCACTTC	2303	GARGGSIVGGVKSACETVIALGATDPELTERAGYALGOKNSKASVILAPTPN
Qy	1945 GCGCAAGTGTCTCACTTCCAAAGCGCTGAGGACACCCCGCTACTCACTTC	2004	HRSGLNGNFEVYSEDPAITACAVAYINGOSQVATIKHFVANSEERQTMSSD
Db	2304 GCGCCCTCTGTCACGAGCTTCCCAGGGCCACGGCCGACCGCGTGTCTCCTCAGG	2363	VDERTLRLYLPFEEAKVAKKAGVAVSYNKLNTTTSNEWLLVILRUEWFGDV
Qy	2005 CCCACCGAGGGCGCCGGCGACGCCCTGAGGAGCTACGGGGTACAGGFACTAC	2054	VMSDWESHHSTAATTINAGLDLMPGPWRGRKLYVAVERSYKAETVYRASRILL
Db	2364 CCGCTCGACGGGACCCCTGAGGACACCCGGCCGCTTCTGCGCTACCGGGCACTGG	2423	LENGFEKKAPOLAHHALDPRALTRQAGAVLNDGVLPLAKSFDQIAYV
Qy	2065 GAGTTCCGACAGGAGCTCAATTCCCTTGGCACCGCCGACGGCTACACCACTTT	2124	PNASARMMGGSGARTIAHAYTSPLEGIRALANSLRHAYGCNNRLDFVSGKT
Db	2424 GCGGACCGGGGGCGCCCGCGCGCTGAGGAGCTACGGGGTACGGFACTGG	2483	VEYFKGRGFESEHVYEVKEFFWIDPSQDLDLADFSAHNTATYFDPGEHIFG
Qy	2125 GCCTTTCAACTCTCCTCGTGTCTCAAGGAGCGCAAGGAGCTGAGGAGCTG	2184	MTRAGLARLWQELVNLDDGWTGEGFIRGEGFIRGEGFIRGEGFIRGEGFIRGEG
Db	2484 GAGTACCCGACGGCGCTGACAGGGACGGCCGACGGCCGACGGCTACCGTG	2533	RUGROEMLIEAVAEIPTVNVVWQTPGPPIEMPWLGVYAVIQLMWNGPQDELGNALDV
Qy	2185 AAGAACCCGGCTCTGCGGGCACAGGTGGCCACGCTCTAGTCAGCCCTCAA	2244	PLKDPYEPGRGIPOTPKLTDQSLPQDQHVRVAFGTVRHRDTRTE
Db	2544 ACCAACACCGGGCGCGCCACCCAGCCCGAGGTGCTGCGCTACTGGAGCCCGTGC	2603	PLPKFGEGLGTYTFWQAPQSLGSTEMQADGLVTVDTNIGRAGSDVQVHSPNA
Qy	2245 GGGCCAGATTAACGGGGCTGAGGAGCTCGAGGAGCTGAGGAGCTGAGGAGCTG	2304	RVVERSPKHLRPAKLLKAPLGAATGAVLKIAPRDLAYDVEAGRFRADAGKYLIVAS
Db	2604 TCCGACGAGCCGCTGCCCTGCTCGCTGGCCAGCGACGGTGACGCCGGCGTCC	2663	AIDRASVSIHLPDHMPE"
Qy	2305 CCGGGAGACG 2316		213. .223
Db	2664 GCGCGCTCAG 2675		/gene="cbg1"
RESULT 8			
ATUCBGI	ATUCBGI	3710 bp	DNA linear BCT 26-APR-1993
LOCUS	A.tumefaciens		beta-D-glucosidase (cbg-1) gene, complete cds.
DEFINITION	A.tumefaciens		
VERSION	M59552.1		GI:142221
KEYWORDS	beta-D-glucosidase.		
SOURCE	A.tumefaciens		
ORGANISM	Agrobacterium tumefaciens		
RHIZOBIAE: Rhizobiaceae, Rhizobium.			
REFERENCE	1 (bases 1 to 3710)		
AUTHORS	Castile, L.A., Smith, K.D. and Morris, R.O.		
TITLE	Cloning and sequencing of an Agrobacterium tumefaciens beta-glucosidase gene involved in modifying a vir-inducing plant signal molecule		
JOURNAL	J. Bacteriol. 174 (5), 1478-1485 (1992)		

QY	434 TCAASCACTTTGTGCAATGATCAGGAGCACGCCATGATAGTGGCAGAGATCGICA	493	Db	1664 CCGCACCGCCGTCACCGGTTGTGGTCAATG-----AGCGCGAAGG	1708
Db	605 TCAAGACTTCTGTCGCCAACGGAGTCGAGATCGAGGGAGCATGTCCTCGATGCG	664	QY	1574 GGCACACCATCGTCCCCGCCACGGCTCCCTCGCGCTCGGGGCTCAAGGTATGACCG	1633
QY	494 CGGAGGGGGCTCCTGGTGAATATCGAGACTCCGTTCCAGATGCTGCGAGACTCC	553	Db	1709 CCAGCCTCTGGGCAATCACATATGTCGCGCTCGCCAGGGAGGAGAAGGGCG	724
Db	665 ATGAGGGGAGCTGGCAATCTATCGCCGCTTCAGGGAGGAGGAGAAGGGCG	724	QY	1634 ACCAGGCGGAATCGAAAGTCTGGTCCCTGCGCAGAGGACGAGCTCATCAT	1693
QY	554 AGCGGGTGTCTCATGACGCCAACATGCAATGCAATGCGCTGTCAGGAGAACCC	613	Db	1766 GCGATGGGGATTGCGGGGTCGAAACGCCGCAACTCGGATATGTCCT	1825
Db	725 GCGTGAAGGCCCTCATGTCCTCTAACACAGCTAACAGGCACTATAGGAGGAATC	784	QY	1654 GCGGGGCTTAACGCCGACTGGAGAGCGAGGGCGCCGACGCCGAGGAGG	1753
QY	614 CTAATATCTGTATGGATCTCGAAAGGATGGGTTGGATGGCTTATATGAGGG	673	Db	1826 TCGTGGGGTGGGGAGGACCCGAGGAGGCTGCGACATGCGCTGCGATGCGCTG	1885
QY	734 CCGGACCTCCACGCCCTCGAGGAGAACACTCAAGTCACGCTCAACGGGAAGCCT	793	QY	1754 CGCGCGCTGGGACATGCGCACGCTGCGGCTGAGGAGGCTGAGATGCC	1813
Db	905 CGGGCCTTGGGGATCGCGGGAAAGCTGTCGGCGAGGATGGGGATGTCCTCG	844	QY	1886 CGGTGCGCAGGAGGAGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGG	1945
QY	794 TTATCAGCTCATGACCGAGGGCTAGGGAGTCTCTAGTGTCTGCAAGAGCTG	853	Db	1814 TCATCGAGCGGCCACCTATGCGCACGCGCCAGCCACCCAAACCGAGCTG	1873
Db	965 AGGCCGAGACGCTGGCGCTTGGGACCGGGTATTCCTCTGCGACGGCTCG	1024	QY	1946 TAATGCAACGGGGTGGTCACATCGAGATGCGATGGCGATGGCGCTG	2005
QY	854 CCTCCGGAGTACGGAGAACGGCCGAGAGACGACTGTCAACAAACCCCGAAGGGCAG	913	QY	1874 AGGCCTGGTACCGGGCACAGAGACGAGACGAGGAGGAGGAGGAGG	1933
Db	1025 CTTTGAAGGGCCGATCTCGCGAACACGCCCTGATCTGGAGATGTCCTGGCC	904	Db	2006 AGATGTGTATCCGGCCAGGAACCTGGCAATGCGCTGCGGAGCTCTTGCG	2065
QY	914 CTCTCTCCGGAAGGTTGGCAACGGGCTGGGACCGGGTATTCCTCTGCGACAGGTC	973	Db	2126 TTACCGACATCGCTGATCTATCTCGGCCAGGACGGCCATCGCGCTACCGGAA	2185
Db	1085 TCAATGCCAACCTGGTGGGGTGTACTCTCTGAGAGATACGGAGTCTGGCC	1144	QY	1934 ACACCCCTCGGCAACGGGCAACTCCATGCGCACGCGCAACTCCATGCGCACG	1993
QY	974 TCGCCCTTGGAGGAGAAGAAGAGGCTGATGCGGCCAACGCCAACGGCCAT	1033	Db	2066 TCGGCCCTGGGCTGCGCTGCGACAGACCGCTTCCCGGAGGAGGAGGAGG	2125
Db	1145 TGGCCAGTCTGCTGCGACAGGAGTCGCGCTATCGGCCCATCGGGCTCGACCG	1204	QY	1994 TTCTCACTTCCGAC-----CGAGCCGGGGCACGGCTGAGCTGCGAGGAC	1873
QY	1034 ACCACGGGGCTCTGGCCACTCAGGCCCTACTACCGAGTCACTCCCTTGACGCC	1093	Db	2185 TCTTGTGGCTATCGCATTACAGTACAGAGATGAGGAGATGCGAGGAGG	2041
Db	1205 TCATGGGGAGGAACGGCGCTACGCCGCGGATGCGCCATTACGGGCACTTGGCA	1264	QY	2102 ACGGCCTGCTCACCCACTTGGCTTTCGAATCTCTCGTCTCACAGGA-----	2156
QY	1094 TCAAGAACGAGCTGAGACGACGCCATCGTACACCGCTCGGGCTACACCACGGT	1153	Db	2245 TCGGCTTGGCTACACCGCTTACCTGGGGTCCCCGCACTATCGGAACGGAATG	2305
Db	1265 TTGGCGGGGCCCT-----GTCCACGGCCATCGGCCATGCGCCTACCGGAA	1103	QY	2042 TCTAATGGGGTACAGGACTTACAGGAGTGGCGCACAGGACCTCAATTCCCTTG	2101
QY	1154 CTCATTCTAGGGAGCGTGTGCTACGCCGAGGGGCTCGGCATGCGCTGGAGGT	1213	Db	2186 TCTTGTGGCTATCGCATTACAGTACAGAGATGAGGAGATGCGAGGAGG	2245
Db	1304 CGGTGGCTGTACACAAACGGCTATCGAGCTCTCGGGCGAGATGGGGAAT	1363	QY	2123 AGCGCTTGGCTTACCCGCTTACCTGGGGTCCCCGCACTATCGGAACGGAATG	2272
QY	1214 TCAACGAGCCCTGTGTACCCCTAACGCCAGAACGACATTGAGGAGCTCTTCA	1273	Db	2366 ACGTCGGTGCAGCTATGCTACTCTCCAA-----TGCAGGGCGACGGCTGG	2422
Db	1364 ACTTCAAAGGGAGCGGGCTTGTAGAGGCCCTCGGTCTACGAGCTGGAGAAGGGC	1423	QY	2273 AGCTCAAGGGCTTGGCAAMGGTGAACCTGAGCCGGGGAGAAGGCC3TG;ATTCG	2332
QY	1274 CGGACATGCACTGGGACTACTACACCCCAAGGGGGAGAACGACGTGGTACGCCA	1333	Db	2423 AGCTGGCTTGGCAAGGCTCAAGCGGGCCGGGGAGGAGGAGGAGGAGG	2479
Db	1424 AATTCCTCTGGCTCTCGTCTCGTCCGGCACCTGTATCGTCGCCATTGCGCA	1483	QY	2393 AAAGGGTACATGAGTTATGTGAGGAGAAGC 2427	
QY	1334 TCGAGGGACGCAACACGCCAGGGACTACCTACGGAGCTGGCTGCAACCCACCAAGGGTCC	1393	Db	2540 ATCGGGCACTACGGCTGAGCTGGCGGCCAGC 2574	
Db	1484 TACGGGACCTCTCGTCTCGGCAACCCGGAGAACCGGTGAACACATCTCGCATGCTG	1543			
QY	1394 GCGCGCAAGGGTAGTAGGAGCACGCTGGCTGCAACCCACCAAGGGTCC	1453			
Db	1544 GCCTTGTGCTGCTGTCGAGCTGGTGTGATGCTATGAGGTTGGAGCA	1603			
QY	1454 CGCGCGATGCGCTCTCGGCTCGGCCACCCGGAGGAGACGGGGCGCATGATCGICA	1513			
Db	1604 AGGTGAGAACTTTGGAACGACGCCAACGGAGCAGGCTGGGGTACGGTTGGG	1663			
QY	1514 AGGGCACACCGTACAAGATCGAGTTGGCTCGCGACCCACCTACACCTCAAGG	1573			
	RESULT 9				
	AB1293760				
	LOCUS	AB1293760			
	DEFINITION	Agaricus bisporus mRNA for putative beta glucosidase (bgl1 gene),			
	Strain	D649.			
	ACCESSION	AJ293760			
	VERSION	AJ293760.1			
	KEYWORDS	beta glucosidase; bgl1 gene.			
	SOURCE	Agaricus bisporus.			





QY	1147	GTTCCTCCCATTTAGGGAGCAGCTACCCGACGCCATGCCCTCACCCGACGCCATGCCCTGG	1206	QY	2209	GCACAGGTGCCAGCTCTAGTCAGTCAGCCCTCCAGGCCAGATTAACCGCCCGTC	2268
Db	80015	TATC-----GTTTCAGCGGCTTAATCGAAACACCCACACCTTCGAGTTTCAAGG	79963	Db	78980	TGGCAGTGGCAAACTATGTCGCGGAGTTGAGAAGCAGCGGTT	78924
QY	1207	AGGCTCTCACCGGACCGCCCTGTTACCCCTAACCGGACGAACTGAGACCTCTTC	1256	QY	2269	AAGGACCTCAAGGGCTTCGCAAGGTCGAACGTGAGAAACCGTGGAG	2314
Db	79962	AAGGAGCTAGCGACCGCT-----GAAGGCTCGAGACCTCTTC	79914	Db	78923	AAAGAGCTGAAGCCCTCTGAAATTGCCCCTGGACCCGGTGA	78878
QY	1267	ACCAAGACGGACATGCACTGGACTACTACACCCAGGACATGAGACGCTCTTC	1326				
Db	79913	CTTGTGTTAGGTTGCCCCGGCCAGGCTTTCGACGCCCCCTTCG-----TTTTCG	79857	RESULT 11			
QY	1327	GCGCACATGGAGGACCGCTACACCGGACGAGACTGACCTACGAGCTGGCTCGTC	1386	AX489445	AX489445		
Db	79856	GCTAGAACTGCAACCATCTCAAGCCTCGCTATCGCTGAGACGCCGCTATCGCTG	79797	DEFINITION	Sequence 6745 from Patent WO20053728.		
QY	1387	GTCTGGGACGCGAACGCTAGTACGACGACGAGCTGCTGAGACGACGAGCTGCTG	1446	ACCESSION	AX489445		
Db	79796	TGGCGCCGGCTTGGCGGCTTATGTCGAGGCGAGCTGTTGTTGAGCTGGGCTTC	79737	VERSION	AX489445.1	GI:22223457	
QY	1447	CAGGCTCCGGCGATGCTCTTCTGGCTCCGCCACCCGGAGAGACGGGCGCATCA	1506	KEYWORDS			
Db	79735	TGGATACCGGTTGACATCTTCTTGAGAAGGGTGCAGTCGCTGAGCTAACG	79677	SOURCE	Candida albicans.		
QY	1507	CTCTCTCAAGGGCACACGTTACACATCTTCTTGAGAAGGGTGCAGTCGCTGAG	1566	ORGANISM	Bukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;		
Db	79676	CTCGAAGCGGCCCTACATAAGGTCGTCGCCAGACGACGTCAAC	79617	REFERENCE	Saccharomycetales; mitosporic Saccharomycetales; Candida.		
QY	1567	CTCAAGGGACACCATGTCGCCAACGGCTTCCCTCGCGCTGCGCTGCAAGTC	1626	AUTHORS	Roemer, T., Jiang, B., Boone, C., Bussey, H. and Ohlsen, K. L.		
Db	79616	CTCTATATCGCCAAATGGTAGGAAATGGC-----AGA	79581	TITLE	Gene disruption methodologies for drug target discovery		
QY	1627	ATTCAGACGACCGCGAACATCGAAAGTCGCTGCCAAGAGCACGCCAGTC	1686	Patent	WO 02053728-A1 (2002-07-11)		
Db	79580	TTTCTGGCGCAACGGAAATGGCGACGCGCCAGCGCGGAGGATG	79521	FEATURES	Eltria Pharmaceuticals, Inc. (US)		
QY	1687	ATCATCTGGCGGCCCTAACGGCAGCTGGAGACCGAGGGCGACGCCGAGGATG	1746	source	Location/Qualifiers		
Db	79520	GTATGCTCTGTCGSCAGACGGCGACTGGATACCGAGGCTCCGATCTCGCGTATT	79461	1. 2601	/organism="Candida albicans"		
QY	1747	AACTCTCCGGCGAGCTGGACGACGCTAACGCGAGCTAACGCGAGGCTCCCAAGGATGAGATCA	1806	BASE COUNT	917 a 334 c 484 g 866 t		
Db	79460	GCACCTCCGGCTCGCGAACGACCGCTGTTGAGCGGTAATTGGCAATCCGACAGC	79401	ORIGIN			
QY	1807	GTCCGCGTCAATGAGACGGCACCCAGGAAATGGCCTGCTGAGGCCAGGCC	1866	Query Match	8.2%; Score 205; DB 6; length 2601;		
Db	79400	ATCGTGTGACTACACCAACCGTGGACCGGTAATGGCTGTTCTCGTGTC	79341	Best Local Similarity	46.7%; Pred. No. 7.6e-16;		
QY	1867	GTCTCCAGGGCTGTCACGGGCAACCTCATGGCGACGCTGCTTT	1926	Matches	1031; Conservative 0; Mismatches 1097; Indels 78; Gaps 9;		
Db	79340	GTACTGCAATGCTGCTATCCGGACAGGGCTGCAACGGGATGCGCTCTC	79281	QY	11 TTGATGTTGAGGCCACTTGTGAGAAGCTCACCTGCGGAGGCTGATCTGCGCTG	70	
QY	1927	GGCGACTACACCCCTCGGCCAACCTGTCCTAGCTTCCCAGCGCTGGAGAAC	1986	Db	14 TTGATGTTGAGTATATTATCTCAATGAGCTTGTAGAAGAAAATTGGTTAGTAGGT	73	
Db	79280	GGCAAGGCCAGCCCTGGCGGACTGGCACACGCTTCCGGCTGGCTGACAC	79221	QY	71 GATCGACTCTGGCACACAAAGGCTCCCAAGGATGAGTCGCTCCCTCTCCCTTA	130	
QY	1987	CCCGGGTCTCTCAACTTCCGACCG-----AGCCGGCGCACCGCTACGG	2034	Db	74 GAATTGATTCTGGCATACATTCATTCATTCTCGATTAACATTCGAATCCAAAGTGGAGATCA	133	
Db	79220	CCACACATCTGAGGATGACCGCTCATCCAGTAAGGACGCCATGTGCATATCA	79161	QY	191 TCCCTTGGCGCACGCTCTGGCTCCACATCACCAACACCTGCGTGAAGAGGAGGA	250	
QY	2035	GAGGAGCTTACGTCGGCTACAGCTACTACGAGTTGCGACAGGAGCTCAATTC	2094	Db	194 TCCCTTGGCGAACCTGGTTGCTGCTACTTGTGATGATTAATTGGTACACGGGA	253	
Db	79160	GAAGGTGTTGCTGCGGTTATCGGCACATACGACGCCACGGCTGCTCCG	79101	QY	251 AGATGATGGCAAGAGGCCATCGCTGAGTGGCGATGATCCGGCGACTATCA	310	
QY	2095	TTGGCCACCGCTCTCACACCACTTGTCTTCCATATCTCCG-----TGCT	2148	Db	254 AATTAATGACATTAATGAGCTTAATTAAATGCGACGATCTGGCTTACGATGA	313	
Db	79100	TTGCTCTGCTGCTGGTTATTCAGCTTGTGCGACTGCGCTAGCTGCTACCA	79041	QY	311 ACATGGCAAGCTCCCTCTGGAGCTGGCTCTGAGTGGATGTTGAGATCGTT	370	
QY	2149	CACAAAGCGCAACCTGACGGCTGCGCCCTCCCTGAAGAACACCGGCTCCG	2208	Db	314 ATATTCAAGGAGGCCATTGGTGGCGGAGGATGATCATTTCTGAAGATCCATT	373	
Db	79040	GACGCTGCCGGCGGGTGAQAGGAGCTGGATGACCAACATCAGGGCGCCAGG	78981	QY	371 TGGGGTTGGAGCTGGCTCATCGGGCATTCAGCAGCTGAGGAGCTGATG	427	
QY				Db	374 TAATCTGGACAAATTGCGCAATTAATTGGTTCAATATGATAAATGAAATGGT	433	
QY				QY	428 CTACGATCAAGGACTTGTGTCATGATCAGAGGAGGCCAGTGTGAGCA	487	
QY				Db	434 CCAGGGTAACATTATGTTGATGATTAATTGAGGATGCTGATGTT	493	
QY				Db	434 CCAGGGTAACATTATGTTGATGATTAATTGAGGATGCTGATGTT	493	
QY				Db	494 TGGTACCAAGAGCATTTGAGAATTATTAGACCAATTAGATACTTAAAG	553	
QY				QY	548 ACTCCAGGGGGTGGCTCATGAGGGTACATGGCATCATGGCTTGCGACCG	607	

Db	554	AAAGTAACTCAATTGTTGAGACAGGGATAATAAAGGTTATGGAGACATGTTCTC	613
QY	608	AGAACCTAAATATCTGTGATGGATGCTTCGGAAAGGATGGGGTTGGATGCCTAATCA	667
QY	614	AAAGTAATTTCCTGCAAAATATATACGTGATGATGGATGGAAAGGACATTA	673
Db	668	TGAGGCACTGTTACGGCACATACAGTACCAAGAAGGATGGGGTTGGATGCCTAATCA	677
Db	674	TATCTGATTTCTGCAAAATATATACGTGATGATGGATGGAAAGGACATTA	733
QY	728	AGATGCCGGAACCTCAGCGCTCCGGAGGGAGAACACTCAAGTCAAGCTCCAAAGGA	787
QY	734	AAATGCCAGGATCTCAATTTCGGTAAATAACAACTTGTGACTAGTATGATTAATCA	793
QY	788	ACCCCTT--TATCCAGCTCATGACCAGCAGGGCTAGGGAGTCTCTCAGTGTCAAG	843
Db	794	AAGAATGCGATCAACATTTGGATGATCGTGTAAAGAAGTGTGTTAAATTAATTA	853
QY	844	-----TATGGTCTGCTCGAGTCAGGAGAACGCCCGAGAGACITCAACAA	898
Db	854	TTCGCCAACAGTCAGTGTGTTGACTGAGATGATGTCAAGTGTGAA	913
QY	899	CCCCCGAACGGCAGCTCTCTCCGGAGGGTGGCACGGGGCATCGTGTGTCAGA	958
QY	914	CTCAAGAACAGAGATATACTACGAAACTGATCAAGAATCAATTGTGAA	973
QY	959	ACGAGAACACCTCTCGCCCTGAGCAAGAG--	1000
Db	974	ATGATATAATTATACCGATGATGAGCTCGTCACTCAGTCAGTCTATTGAT	1033
QY	1001	TGATTCGCGGCCAACGCCAGAGGCCATACCCAGGGAGSCTCTCCGACTCA	1060
Db	1034	CAATAATTGTCTCTAACGAAATGCGCTTATTCAGTGTGGTCACCCATTTAC	1093
QY	1061	GGGCCATRACTGGCAGTCACCTCCCTGACGCCCTCGACAGCAGCTCGAGAC	1120
Db	1094	CAGCTTATTACCACTACTCTTAAATGCCATTTGTGAGAAATTATCTGTATTC	1153
QY	1121	CGTACACCGTGGCGCTAACCCACGGTCT-----C	1153
Db	1154	AATTGATATTACATCACACTCAAAATACATATTGTTGCAAAAGCATATAATTAC	1213
QY	1154	CCATCTAGGGAGCAGTGTGTCAGCGCCGAGGGCTCCGGCATGCGTGGAGGT	1213
Db	1214	CAGAATTAGTGTCCACAGTGTCAATTCACAACTCTGTAAACCCGGTTTCATGAAT	1273
QY	1214	TCAA-----CGAGCCCTGGTACCCCTAACGCCGCCACATGAGGACTCTC	1264
Db	1274	TTTATAAAAGCCAAATCCGTCCTAATGAAATTGAGAATTATTGATGAATAGATA	1333
QY	1265	TCACCGAGGGACATGCGACCTGGTGGACTACTACCAACCCAAAGGGGCAACAGGTG	1323
Db	1334	CAGAAATATGTATATTAGTGTGATTATTCATAAGATAATCCCTCTAATAGGT	1393
QY	1324	-----TACGGCAGACAGGGAGGGACGACCGCCGAGCGGGACTGCACTACGAGCTCG	1378
Db	1394	TATATTATGTGATTGTTGATGATTCAACCATCAAAACCCACATATTGATTC	1453
QY	1379	GCCTGTCGCTCGCAGCGCAANGGGCACGTAGACGACGAGCTGCGTGTGACACCG	1438
Db	1454	GTTAACTGTTACGACACATTATTATGATGATAACTTGTGTTGATAATA	1513
QY	1439	CCACCAAGCAGTGGCCGGGATGCTCTCGGCTCGGCCACCGCGAGGAGGGCC	1498
Db	1514	AAACCAACAACTTAAGGGAGTTCTTGTAACTTGTGAAACCTTGTGAAAGCAGGAT	1573
QY	1499	GCATCACTCGTCAGGGACAGCTACAGTCAGTCAGTACAGTGTGAGTGGTCCACCA	1558
Db	1574	CAATGCAATTACATCAAGGTAAACTTATAATCATGTGTGAAATATGGTCCGGACCA	1633
QY	1559	CCTACACCCCTAACGGCGACACCATGTCGGGACACGCTCCCGGACGGGGCT	1618
Notes: 15			
COMMENT			

Streptomyces coelicolor sequencing at The Sanger Centre is funded by the BBRC.

Details of *S. coelicolor* sequencing at the Sanger Centre are available on the World Wide Web.

(URL: <http://www.sanger.ac.uk/Projects/S-coelicolor/>) CDS are

numbered using the following System eg SCB7\_01c SC (S

coelicolor), 7B7 (cosmid name), .01 (first CDS); c (complementary strand).

The more significant matches with motifs in the PROSITE database are also included but some of these may be fortuitous. The length in codons is given for each CDS.

Usually the highest scoring match found by fasta -o is given for CDS which show significant similarity to other CDS in the database.

The position of possible ribosome binding site sequences are given where these have been used to deduce the initiation codon. Gene prediction is based on positional base preference in codons using a specially developed Hidden Markov Model (Krogh et al., Nucleic Acids Research, 22(22):4768-4778(1994)) and the FramePlot program of Bibb et al., Gene 30:157-66(1984) as implemented at <http://www.nih.gov/>

/cgi-bin/frameplot.pl. CAUTION: We may not have predicted the correct initiation codon. Where possible we choose an initiation codon (atg, gtc, ttg or (att)) which is preceded by an upstream ribosome binding site sequence (optimally 5-13bp before the initiation codon). If this cannot be identified we choose the most upstream initiation codon.

IMPORTANT: This sequence MAY NOT be the entire insert of the sequenced clone. It may be shorter because we only sequence overlapping sections once, or longer, because we arrange for a small overlap between neighbouring submissions. Cosmid 8A6 lies between 3F9 and 1F2 on the *Asel*-A genomic restriction fragment.

FEATURES Location/Qualifiers

source 1. 3745

/organism="Streptomyces coelicolor A3(2)"

/strain="A3(2)"/

/db\_xref="txon:100226"

/clone="cosmid 8A6"

2. 107

/note="nominal overlap with StB96 from 1 to 106

nominal overlap with cosmid 3F9 from 19725 to 19830"

567..711

/note="hairpin loop with 23 bp stem"

gene complement(730..2160)

/gene="SC06581"

complement(730..2160)

/note="SC8A6\_02c, probable transmembrane transport

protein, len: 476 aa; some similarity to e.g. TR:051330

(EMBL:Id007\_5) oxalobacter formig formicenes (418 aa), fasta scores: opt: 226

z-score: 230.1 E(): 1..4e-05, 24.1% identity in 436 aa

overlap",

CDS /codon\_start\_1

/transl\_table=11

/product="putative transmembrane transport protein"

/protein\_id="CA1974\_1"

/db\_xref="GI:3288602"

/db\_xref="SPREMBL:087836"

/translation="MGGCAGRSFDRAGDAIDMTTDTMTRVAAREVDRNGFVRYGE

SDIDLMGRKKWVILPNWGMGMISSAYAFFASAEDTIEHAODWHSSTIYMMTAWF

COAQAFFPAGRGRLESKGKPARHMMGSAAGTIGYSLAFAPHVSLAFIGGVFSENG

AGMFTATCVNMWKGWPERPGKTFGGTGGFAGSVVFFGKMDTSNFRWLYSA

GVFLAAWVATAGYFPRDPKNNPAVADPLNPADPARRSLEKNPRAVOKSPMRW

KTRGVAMWFCFLACTSGSYNIFGTAFOFDIGERGAFAAGVVAAMSRSKAGVTRGRTI

GWLSDLYGRKQCLWVATGILQAFNLSKLNFLVLSAISCGGAIIFPRA

ALIADYFGENNNTAYGNYGAVYSSKLVSGIAGGMSVVVWGGNGATSLAGS-SIFAGF

VALSLRPPGCRPKGRKRVENPOPLGEG"

complement(2233..4377)

/gene="SC06582"

complement(2233..4377)

/gene="SC06583"

hypothetical proteins including several from each of the

archaea on Archaeoglobus fulgidus TR:028341 (EMBL:AE000970)

conserved hypothetical protein ARI938(673 aa), fasta

scores: opt: 960 z-score: 1289.8 E(): 0, 35..6% identity in

693 aa overl ap and *Pyrococcus horikoshii* D108624

(EMBL:AB009518) hypot heticl protein FHCV016 (460 aa),

identity in 461 aa overlap. Contains PS00216 Sugar

transport proteins signature 1. Also similar to *S. coelicolor* hypothetical protein SC9B10\_09 (E(): 0..44..3%

33..9% identity in 375 aa overlap"

/codon\_start=1

/transl\_table=11

/product="conserved hypothetical protein"

/protein\_id="CA1975\_1"

/db\_xref="GI:3288603"

/db\_xref="SPREMBL:087837"

/translation="MADDLARVGLDVAWRGEGRTALTAPECKVLADAYGIAVPGEE

RAYDSARLEGVOIETLLOLOGOEVYGVATDPFGKVVAFGLGGVWFLWKLDFRLA

FVDADEALMSDISTRALRIVLVRGQGQDWRVVAQDVRVSEWLADPEVADWV

VIATTEGAVNLLDGGFEGETHYVPKDYLGRKAVSVTDVPGENDVVAVAPKVAAC

GNSVNRNLVDDGEGFETHYVPKDYLGRKAVSVTDVPGENDVVAVAPKVAAC

LEEVGRKTFPMLVPLPSGDRARAKTFFPGAGNPNVDTGGEPSPSYEATIRGLDPRH

ATFCIPYDVGGVGVALTSOGGIGMAILGARTTIGVSAVUGLNKSQDDEDDLW

GDEPTECTAHLDELDKDRFARVAVARAVPRKPVWVLAAGRITAGKAGKASINGLAL

GDDAVYDILRQGVTAGCINLEYARVPLVPAWPECDNNVWITGGAGVWFLWKLDFRLA

VTDNGLSLMLPPLDPRARAKTFFPGAGNPNVDTGGEPSPSYEATIRGLDPRH

ALVIGYWHHTVTPVAVFAELTARYVAEERBERGKPVWVSLAGDVVEEACOYLFERG

VVAVYTTKPVYAVGAKFWRWAAQAGLIGGGS"

complement(3340..3387)

/gene="SC06582"

/note="PS00216 Sugar transport proteins signature 1"

complement(4382..5614)

/gene="SC06583"

complement(4382..5614)

/gene="SC06583"

/note="SC06583"



QY	1737	CGCGAGCATGAACTCCCGGCGCTGCGACCGCTCATGCGACCTGGCGCCGGAA	1796	Db	1151	CCCTGGCAGACCTTCGAGACTCCATGCCGAGACTCATGGCGAGGCGCATGGCGCCGG	1210
Db	29404	TAGGACCTGGCCCTGCAGCCACCCAGACGGCTGGCCCTCGCCGCA	29463	QY	266	AGGCATCGCTAAGGATGTCGCTATGATCTCGGCCGACTATCACATCAAGCTCC	325
QY	1797	CCCAACACCGTGTGCGTATGAGACGGCACCCAGGAGATGCTGGCG	1856	Db	1211	ACGGACCGCGCTGGGCCAGACATGGTCTGGGCCGATGACAACAACTGGGTC	1270
Db	29464	CCCGCCACCGTGTGCGTACAGGAGCTGGGAGCTGGGAGCA	29523	QY	326	CTCTCGGTGACGTGGCTTCTGAGTGTGAGATGCTCTGGGGCTGGG	385
QY	1857	CAACCCACCGTGTGCGTACAGGAGCTGGGAGCTGGGAGCA	1916	Db	1271	CACACGGCGGAACTACAGAGACTTCAGAGCTTCAGAGGAGCTGGGAGCA	1330
Db	29524	GGCGGGTGGGTGTGCTGGTTCGGACAGAGGGCGCTGGACTGGCGA	29583	QY	386	CTGGGGCTCATCGGGCATTCAGAGCTGGACTGGACTGGCTAGATCAGACCTT	445
QY	1917	CGTGTCTTGGGACTTACACCCCTCGGGAGCTGRCCTAGCTTCCCAGCCT	1976	Db	1331	CGTCGCCAGTCAGGGATCCAGGGTCTGGGCTGAGGACACCAGGGCA	1390
Db	29584	CGTGTCTTGGGACTTACACCCCTCGGGAGCTGRCCTAGCTTCCCCT	29640	QY	506	TCTGTAATCTACGACTCCGTCAGGGTCTGGGCTGAGTGGCTAGACGA	565
QY	1977	GCAGACACACCGGTTCTCACTTCAGGAGTACTACAGCTGGCTGGCGA	2036	Db	1451	TCGGCAATGTCAGGGACAGGCGATGATGTTGCGACAAGGAGCTGGGTC	505
Db	29641	CTCTGGGACGCCCCGGTACCGGACCCGGCCGACGGCGGACCTACCTGGCG	29700	QY	626	ATGGATGCTTCAATCTCGTGTCTCGTGTCTCGTGTCTCGTGTCTCG	625
QY	2037	GGACCTCTAGTCGGTACAGGAGTACTACAGCTGGTGTGCGACAAGGAGCTGGCG	2096	Db	1391	CGGCCAACACACAGGAGAACACCGCTTCAGCTAACGGTCAACGCGTCAACGAGAC	1450
Db	29701	GGACTGCGACCTCGTACCGGACCCGGCTGGCATCGGACAGCTGGCTACTGCTT	29760	QY	566	TCTGACGGGACAATCGCATCGCATCGGCTGCGGCAAGGAGACCCCTAATCTCG	1567
QY	2097	TGGCACGGCTGCTTACACCTT-----TTCGTTTCAATCTCGTGTCT	2147	Db	1508	TCTGTAATCTACGACTCCGTCAGGGTCTGGGCTGAGTGGCTGCTGGCTG	1507
Db	29761	CGGACACGGCTCGCTACAGACTGGCTGGAGGAGCTGGCTGGCGTAC	29820	QY	626	ATGGATGCTTCAATCTCGTGTCTCGTGTCTCGTGTCTCGTGTCTCG	625
QY	2148	TCTACAGACGGCAGCTGAGCTGAGCTGGCTCTCGTGAAGAACCGGCTCCGG	2207	Db	1568	ACACGTCTGCCACCGTGGCTCAGGGCTGGCTGACGGTCCGG	1627
Db	29821	CGGGGGGGGCGACGCGCTACCGTGTGGCTGGGGTGGCGAACACCGGTGCGAGGG	29880	QY	686	CATACAGTACACAGAGCTGGCTGAGGAGCTGGGCTGAGGAGCTGGCTG	745
QY	2208	CGCACAGGTGGCTTCCAGTGTGCTGAGGAGCTGGCTGGGGTGTGAC	2267	Db	1628	C-----CCGGCGCAGGACCCATACCAAGGGCTGACCCAGGAGAT-----GGGGTGTG	685
Db	29881	CGGGGGGGGTCTTCCAGGCTGAGGAGCTGGCTGGGGGGGGGGGGAGCGTGTGGCG	29937	QY	746	GCTTCGAGGAGAACACTCAAGTCACTGTCAGCTTCACGGAAAGCCCTTATCCAGG	805
QY	2268	CAAGGAGCTCAAGGCTTCCAGGAGCTGGCTGGGGGGGGGGGGAGCGTGTGGCG	2327	Db	1681	CTTCGGGGACATCCCGGGGAGCCCTGGCGCGCCAGGAGCTGGCTG	1740
Db	29938	GGTGTGGCTCCGGGTACAGGGCTGGGGGGGGGGGGAGCGTGTGGCG	29997	QY	806	TGACCGAGGGCTAGGGAGTCTCTGGTCTCTGGGAGTGTGCTGGCTGGAGTGTG	865
QY	2328	AACTGAGGAGCAGGAGACTACGAGCTGGCTGTATTTGTGAGAAGC	2374	Db	1741	CTGA-----AGCGCGCTCTGACGGCAGCGTCCGGCG	1780
Db	29998	GTTGCGTCCGGGGCCCTGGCCACTGGTGGTGGAGGAGC	30044	QY	866	CGGAGAACGGCCGAGAGGAGCTGACACACCCGAAAGGGAGCTCCCGA	925
RESULT 13							
AR1732							
LOCUS	AR1732	3241 bp	DNA	linear	PAT	17-DEC-2001	
DEFINITION	Sequence 11 from patent US 6303767.						
ACCESSION	AR1732						
VERSION	AR1732.1						
KEYWORDS							
SOURCE							
ORGANISM	Unknown.						
REFERENCE	Unclassified.						
1	(bases 1 to 3241)						
AUTHORS	Betlach, M.C. and McDaniel, R.						
TITLE	Nucleic acids encoding harbonolide polyketide synthase enzymes from streptomyces harbonensis						
JOURNAL	Patent: US 6303767-A 11-16-OCT-2001;						
FEATURES	Location/Qualifiers						
source	1..3241						
BASE COUNT	520 a 1174 C 1125 g 422 t						
ORIGIN							
Query Match	7.4%	Score 184.4;	DB 6;	Length 3241;			
Best Local Similarity	48.2%	Pred. No. 2.1e-13;	Mismatches 1028;	Indels 150;	Gaps 12;		
Matches	Conservative	0;	956;				
QY	206	CGCTGGTCCACATGACCAACAACTCTGCGAGAGGGAGGTAGATGATGCGAAC	265	Db	2126	CTGGTACCCATAACGGCGCACATTGAGAGCTTCCTCACAAAGGGAGATGAC	1285
				Db	2140	ATCCGGGGCGCGCAGTCAGC-----CGGGCTTCAC	2173
				QY	1106	TGAGAGACGGCCACGTCACCTGGGGCTACACCCAGGGCACCCAGGCTGGCG	1165
				Db	2020	CTGGGAGGCCAGTCACCTGGGGCTACACCCAGGGCACCCAGGCTGGCG	2079
				QY	1106	AGCAGTGCTCACGCCAGGGCTCCGGGAGCTGGGGCTTCACGAGCCC	1225
				Db	2080	CGCCGGGGCGGGGGAGGGTACGAGAGCTGGGGCTGGCG	2139
				QY	1286	TGGTGGACTACTACACCCAGGGCGAGACGGTGTGGAGGAGCTTCGGAC	1345
				Db	2174	AGGCCACAGCTGGAGGGCGAGGGGGCTGAGCGGAGCTACCGTGC	2233

DEFINITION	Streptomyces narbonensis desosamine biosynthetic gene cluster, partial sequence.
ACCESSION	AF521878
VERSION	AF521878.1
KEYWORDS	
SOURCE	Streptomyces narbonensis.
ORGANISM	Streptomyces narbonensis.
BACTERIA; Actinobacteria; Actinomycetidae; Actinomycetales;	
STREPTOMYCETACEAE; Streptomyces.	
REFERENCE	1 (bases 1 to 17665)
AUTHORS	Bate, N. and Cundliffe, E.
TITLE	The desosamine biosynthetic cluster of Streptomyces narbonensis, producer of narbomycin
JOURNAL	University Road, Leicester LE1 7RH, UK
FEATURES	JOURNAL/Qualifiers
source	Unpublished
REFERENCE	2 (bases 1 to 17665)
AUTHORS	Bate, N. and Cundliffe, E.
TITLE	Direct submission (14-JUN-2002) Biochemistry, Leicester University
JOURNAL	University Road, Leicester LE1 7RH, UK
gene	
CDS	
QY	1346 AGACCGCGGAGGACTGCACCTACGAGCTCGGCCCTGCTCTGGGCACGGCAAGG 1405
Db	1466 TCTTCGCTCCGCCACCCGGGAGAGGGCGGCCGATCACTCGTCAGGCAACGGTGCAG- 2292
QY	1406 CGTACGTAGACCAACCTCGTGTGTCACACAAAGCCACCAAGCCACCAAGGGCAAGG 1465
Db	2293 ----- CTCGGCAGCCAC 2305
QY	1526 ACAAGTTCAGATCGACTTCGCTCGCACCCACCTACACCTCAAGGGCACACATCG 1585
Db	2366 GCACGCCAACAGT----- CACATCTGGGCTTCGCGATGA 2401
QY	1586 TCCCCGCCAGGGCTCCTCCCGTGGGGCTGCGAGGAGCTGCAAGGTCATGAGGACAGGG 1645
Db	2402 GCGCCACGCCGCTCTCGTGGCTGGTGGAGCGAGGGAGGAGGAGGAGGAGGAGG 2461
QY	1646 TCGAAAGTCTGCGCCCTGCCAACAGGAGACGAGGAGGAGGAGGAGGAGGAGGAGG 1705
Db	2462 TCGCGAGGGCGTGGAGTCGCGCGAGGCGCCGAGGGCATCGTGTGGCG----- 2514
QY	1706 AGGCCACTGGAGACGGAGCGGCCAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1765
Db	2515 --TAGACGAGGCCACCGAGGGCGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2572
QY	1766 ACCAGCTCATTCGCCGAGTCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1825
Db	2573 ACAAGCTGATCTGGGGGTGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2632
QY	1826 GCACCCCGAGGAGATGCCATGGCCGACGGAGGAGGAGGAGGAGGAGGAGGAGG 2632
Db	2633 GTTCTGGTGGCTGTGAGGCCGGCTGGCTCCAGACCCGCGCGCTGTGAGGATGTTG 2692
QY	1886 GGGCCACGAGACGGCAACTCCATGCCAACGGTGTCTTGGCACTAACCCCG 1945
Db	2693 CGGGCAGGGGGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2752
QY	1946 GCAAGGTGCTCTCACTCTC-----CCAGCGCTGAGACACCCCGCGTCTGAGGATG 1999
Db	2753 GCAACCTCAGCAGACCTCCGGCCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2812
QY	2000 ACTTCGCGACCGAGGCCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2056
Db	2813 ACCGCTACCGGGCTGAGAACACCAGCAGAGCTACAGCGAGGGCATCCAGCTCGCGTAC 2872
QY	2057 GGTACTACGAGTTGCCGAGCAAGGAGCTCAATTTCCTTGGCACCGCCCTGCTTACA 2116
Db	2873 GCTGCTTCTGCAAGCAAGGAGAGCTCAAGGCCGCTGTGAGGCCAGCTACGTC 2932
QY	2117 CCACITTCGCTTCCAATCTCTCGTGTCTCACAGAGAC-----GGCAAGCTGAGG 2173
Db	2933 CCTCGTTCAAGCAGGAGGAGGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2992
QY	2174 CCCTCTCCGGTGAAGAACACGGGCTCGTGCCTCCGGCGCAAGGGTGCCAGCTACGTC 2233
Db	2993 CGGTCACGCTGGCAACAGGGGGCAGGGGGCAGGGGGCTGCTCAGGCTATCTG 3052
QY	2234 AGCCGCTTCAAGCAGGAGGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2293
Db	3053 GCGGGAGGCCGAGGTGAGCTCGCAGGGAGGAGGAGGAGGAGGAGGAGGAGG 3112
QY	2294 TCGACTCGAGCCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2327
Db	3113 TCGCGCTCGGGGGGAGTCGAGACGGTGAC 3146
RESULT	14
LOCUS	AF521878
FEATURES	gene
source	17665 bp DNA linear BCT 29-JUL-2002
gene	
CDS	
QY	1346 AGACCGCGGAGGACTGCACCTACGAGCTCGGCCCTGCTCTGGGCACGGCAAGG 1405
Db	1466 TCTTCGCTCCGCCACCCGGGAGAGGGCGGCCGATCACTCGTCAGGCAACGGTGCAG- 2292
QY	1406 CGTACGTAGACCAACCTCGTGTGTCACACAAAGCCACCAAGCCACCAAGGGCAAGG 1465
Db	2293 ----- CTCGGCAGCCAC 2305
QY	1526 ACAAGTTCAGATCGACTTCGCTCGCACCCACCTACACCTCAAGGGCACACATCG 1585
Db	2366 GCACGCCAACAGT----- CACATCTGGGCTTCGCGATGA 2401
QY	1586 TCCCCGCCAGGGCTCCTCCCGTGGGGCTGCGAGGAGCTGCAAGGTCATGAGGACAGGG 1645
Db	2402 GCGCCACGCCGCTCTCGTGGCTGGTGGAGCGAGGGAGGAGGAGGAGGAGGAGG 2461
QY	1646 TCGAAAGTCTGCGCCCTGCCAACAGGAGACGAGGAGGAGGAGGAGGAGGAGGAGG 1705
Db	2462 TCGCGAGGGCGTGGAGTCGCGCGAGGCGCCGAGGGCATCGTGTGGCG----- 2514
QY	1706 AGGCCACTGGAGACGGAGCGGCCAGGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1765
Db	2515 --TAGACGAGGCCACCGAGGGCGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2572
QY	1766 ACCAGCTCATTCGCCGAGTCGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1825
Db	2573 ACAAGCTGATCTGGGGGTGGTGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2632
QY	1826 GCACCCCGAGGAGATGCCATGGCCGACGGAGGAGGAGGAGGAGGAGGAGGAGG 2632
Db	2633 GTTCTGGTGGCTGTGAGGCCGGCTGGCTCCAGACCCGCGCGCTGTGAGGATGTTG 2692
QY	1886 GGGCCACGAGACGGCAACTCCATGCCAACGGTGTCTTGGCACTAACCCCG 1945
Db	2693 CGGGCAGGGGGCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2752
QY	1946 GCAAGGTGCTCTCACTCTC-----CCAGCGCTGAGACACCCCGCGTCTGAGG 1999
Db	2753 GCAACCTCAGCAGACCTCCGGCCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2812
QY	2000 ACTTCGCGACCGAGGCCGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2056
Db	2813 ACCGCTACCGGGCTGAGAACACCAGCAGAGCTACAGCGAGGGCATCCAGCTCGCGTAC 2872
QY	2057 GGTACTACGAGTTGCCGAGCAAGGAGCTCAATTTCCTTGGCACCGCCCTGCTTACA 2116
Db	2873 GCTGCTTCTGCAAGCAAGGAGAGCTCAAGGCCGCTGTGAGGCCAGCTACGTC 2932
QY	2117 CCACITTCGCTTCCAATCTCTCGTGTCTCACAGAGAC-----GGCAAGCTGAGG 2173
Db	2933 CCTCGTTCAAGCAGGAGGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2992
QY	2174 CCCTCTCCGGTGAAGAACACGGGCTCGTGCCTCCGGCGCAAGGGTGCCAGCTACGTC 2233
Db	2993 CGGTCACGCTGGCAACAGGGGGCAGGGGGCTGCTCAGGCTATCTG 3052
QY	2234 AGCCGCTTCAAGCAGGAGGCCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2293
Db	3053 GCGGGAGGCCGAGGTGAGCTCGCAGGGAGGAGGAGGAGGAGGAGGAGGAGG 3112
QY	2294 TCGACTCGAGCCGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2327
Db	3113 TCGCGCTCGGGGGAGTCGAGACGGTGAC 3146
RESULT	14
LOCUS	AF521878
FEATURES	gene
source	17665 bp DNA linear BCT 29-JUL-2002



QY	446	TGTGCAATGTCAGCAGGAGCACGCCATGATGAGTGGAGAGATCGTCACGGGGGCTC	505	QY	1526	ACAGAGTCAGATCGACTTCGCTCCACCCACCTACACCCCAAGGGGACACCATCG	1585
Db	5530	CGGCCAACACCAACCCAGGAGAACACCGCTTCAGCGGACGGACGACGC	5589	Db	6505	GCACGACAACT-----CACGATTCGGCTTGGATGA	6540
QY	506	TCCGTGAATCTAGCACTCCCGTTCAGATGCTAAGGCGACTTCAGCTCAGCGGACGG	555	QY	1516	TCCCCGCCCCGCTCCCTCCCGTGGCTGAGCGGGCTGGCTGAGGAGTCATGACGACGAGCGAA	1645
Db	5590	TCCGGAGACATCGAGTTCGGGCTTCGAG-----GGGRCCTCGAAGGGCGGGCTCCT	5546	Db	6541	GGCCACGCCCTCCGCTGAGCTGGCTGAGCGGGAGCGACGCCAGCGAACGAG	6600
QY	566	TCAAGACGGGTTACATGGCACTAATGGCTGTCAGGAGACCCATAATATCTG	625	QY	1646	TGAAAGTCCTCGCCCTCCCAAGGAGCACGACGGCTCATCATCGCGGGCTTA	1705
Db	5647	TCACTGTCGCTTAACCGCTCAAGGGCTAACGGCAAGCGACGCTCGGGCACGAGGTGCTCA	5706	Db	6601	TCGGAGGGCTGGTGGCTCGCGGAAGGGCTACGGCCATGTTGCT	6653
QY	626	ATGGGATCCTCGAAAGGAATGGGTTGGGATGGCTTAATCTATGAGCAGTGGTACGGCA	685	QY	1706	AGGCCAGCTGGAGACGGGGCGCAGCGCAGCTGAAGCTGCTCCGGCTGCTG	1765
Db	5707	ACACAGTCGCTGGCACCGCAGTGGGCTTCAGGCGCTGGTATGCTCGACTGGCTGCCA	5766	Db	6654	-TACGAGCGACCGAGGGCTGACGCTGCGTCAACTGCTGGGTAACGCGAG	6711
QY	686	CATACAGTACAGAGAGGCCGTCGGCAAGCTCGACCTCGAGATGCGGACCTC	745	QY	1766	ACCGAGTCATGCGACGCTGGCCGCCGCCAACACCGCGCTGAGGACGG	1825
Db	5767	C-----CCGGGAGCGGACGGCCATCACCAAGGGCTGACGGAGACAT-----GGCGTCG	5819	Db	6712	ACAACTGATCTCGAGGGTGGCCACGCTGACGAGCTGCTGACACCG	6771
QY	746	GCCTCCGAGGGAAACACTCACTCACTGTCACGCTCACGAAACGCCCTATCCACGTC	805	QY	1826	GCACCCCGAGGAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1885
Db	5820	CTCCCCGGCGCACATCCGGCGGGCAGCCTCGCGCGCGCAGCTGCTGCG	5879	Db	6772	GTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	6831
QY	806	TGACCAAGAGGGCTAGGAAGTCTCTCAGTTCGTCAGAAGAATGCTCGCTCGAGTGA	865	QY	1886	GGGGACACGAGACGGCAACTCTATGGACGCTGCTGCTTGGGACTACACCCCTCG	1945
Db	5880	CTGA-----AGCACGGCTCTGACGGCACGGTCCCGAGGG	5919	Db	6832	CGGGCAGGGCGCCAGCGGCTCCCGAGGG	6891
QY	866	GGGAGAACGGCCGAGACCACTGTCACACACACCCCGAACGGCACCTCTCTCG	925	QY	1946	GCAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1999
Db	5920	CGTGTAGCGCGTCGGGGAGCGCTACGTCACAGATGGACAGTCGCTCG	5979	Db	6892	GCAGCTCACGAGAGCTTCGGCCACGGACCATACAGGGAG	6951
QY	926	AGGTTGTCGCAACGAGSAGCATGTCGTCGTCGAAAGAACGACGACGTCG	985	QY	2000	ACTTCGCGACCGAGGCCAGGCGCAGCGCAGCGCTGCTGCTGCTGCTG	2056
Db	5980	CGACTCGGGCCGCCGAGCGCTGACAAAGGGGGGCCAGGGCTGACAGG	6039	Db	6952	ACCCSTACCGGGCTCGACAAACAGGAGACCTACAGCAGGGCATCG	7011
QY	986	AGAGAGAAGAGAGACATGTCGCCCCAACGCCAACGGCCACATACAGGGAG	1045	QY	2057	GTTACTACAGTGTGGCAGCAAGGACCTCAATTCCCTTGCCACGCCGCTGCT	2116
Db	6040	TGCGCGAGAACGGCGCTCTGCTCGCCAGGGCGCTCCGGCTGGCG	6099	Db	7012	GCTGGTGGACAAGGAGAACGTCAGCTGCTGCTGCTGCTGCTGCTG	7071
QY	1046	GCTCTGCCGACTCTGGCTACTACGGCTACTCCCTTGACGGCTCAGCAGCAG	1105	QY	2117	CCACTTTGCTTTCGATCTCTCGCTCAGGAC-----GGCAACCTGAGCT	2173
Db	6100	ACGGGGCACGASCATCGCCGTCATCGCCCAACGGCGTCAC-----CCCAAGG	6158	Db	7072	CCCTGTCACGAGAGGCCGCGCCAGGCTGCTGCTGCTGCTGCTG	7131
QY	1106	TGAGACGCGCCATCGTACACCGTCGCGCCATACCCCGTCTCCATTCAGGG	1165	QY	2174	CCCTCTGGAGAGACGACGCCGCTCGTGCCTGGCCACAGGTGCTCCAGGCTACGTC	2233
Db	6159	CTGGCAGCGCCACGTCGTCGCGACTCGGGGGGCCGCTCGACACCAAGGG	6218	Db	7132	CGTCAGGGTGGCGACAGGGCGAGGGCGCCAGGGCTGCTCCAGGCTAC	7191
QY	1166	AGCACTGCTCACGCCAACGGCTCGGGCTGGGCTTCACGAGCCCC	1225	QY	2234	AGCCCTCACGAGGCCAACGAGTAACGCCGCTAACGCTGCTGCTG	2293
Db	6219	GCGGGGGGGGGTGCGACGCTGACGCTACGAGACGGGGAGACCTTGGACG	6278	Db	7192	GGGGACCCGGAGGTCACGCCGCTCCGGAGGGAGAGCTGCTGGCTAC	7251
QY	1226	CTGGTACCCCTAACCGCCAGCACATGACGAGCTCTCTCACCAGACGGACATG	1285	QY	2294	TGAACTGCAACGCCGAGAACGAGGGCTGTAC	2327
Db	6279	ATCCCCGGGGCGCAGCTACGCT-----CGGGCTCAAC	6312	Db	7252	TCGGCTGGGGGGAGTGAAGAGACGGTGAC	7285
QY	1286	TGTTGGACTACTACCAACCAAGGGCCAGAACGCTGGTACGCCACATGAGGG	1345	RESULT	15		
Db	6313	AGGGCACAGCTGGAGGCCAACGGGGCGCTGAGGGCACGCTGACGCC	6372	SEY14327	C		
QY	1345	ACGGGGCACAGGAGCTGACTCTAGAGCTGGCTCGGCTCGGGCACGCCAAGG	1405	LOCUS	SEY14327		
Db	6373	CGCCGAGGGAGTACCGCATCTGGCTCAAGGCCACCGGTGCTACGGACGG	6431	DEFINITION	Saccharopolyspora erythraea	2760 bp	DNA
QY	1406	CCTACGTTAGACCAACGCTGGTCAAGGCCACCAAGCAACGAGTCCCGG	1465	ACCESSION	2760 bp	linear	BCT 11-JUN-1998
Db	6432	-----CGGGAGCCACA	6444	VERSION	Y14327		
QY	1466	TCTTCGCGCTGCCAACCGCGAGGAGACGGCGCATCAATCTGTCAGGGACACGT	1525	KEYWORDS	Y14327-1	G1:2982610	
-D0	6445	CCATCGAGGGCGGTCAAGTCACGCCAACGGCGCTCTCAAGCTGACCAAG	6504	SOURCE	beta glucosidase; eryBI gene; eryBIII gene.		
REFERENCE			Saccharopolyspora erythraea.				
AUTHORS			Bacteria; Firmicutes; Actinobacteria; Actinobacteridae; Actinomycetales; Pseudonocardineae; Pseudonocardiaceae; Saccharopolyspora; 1 (bases 1 to 2760)				
CORTES, J. and Leadlay, P.F.			Gaissler, S., Bohm, G.A., Dounith, M.C., Dhillon, N.,				

TITLE Analysis of eryB1, eryB1II and eryB1II from the erythromycin biosynthetic gene cluster in *Saccharopolyspora erythraea*

JOURNAL Mol. Gen. Genet. 258 (1-2), 78-88 (1998)

MEDLINE 98277631

PUBLMED 9613575

REFERENCE 2 (bases 1 to 2760)

AUTHORS Leadlay, P.F.

TITLE Direct Submission

JOURNAL Submitted (25-JUL-1997) P.F. Leadlay, Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1QW, UK

COMMENT Updated by [3]

FEATURES Overlapping sequences: X00379, M11200.

source Location/Qualifiers

gene

/organism="Saccharopolyspora erythraea"

/strain="NRRL2338"

/db\_xref="Taxon:1836"

1. . 2760

gene "ermE"

<1. . 45

/gene="ermE"

CDS /codon\_start=1

/transl\_table=11

/protein\_id=CAA74701.1

/db\_xref="GI:2763398"

/translation="TGGPGRGRGGPQR"

complement(100..2526)

/gene="eryB1"

complement(100..2526)

/gene="eryB1"

/codon\_start=1

/transl\_table=11

/product="beta-D-glucosidase"

/protein\_id="CAH74702.1"

/db\_xref="GI:2982611"

/db\_xref="SPREMBL:070021"

/transl\_table="MIGGERKRVIRVIAPIPLAIVPLVLAASPVSPVRSRORVDELIGQ

TIDEKUSFVVDYDNERKDPPLAKLWFLPVGVLPIKRGTDGAGTTHOPATIMPAVA

IASAFDRDLAHEYGLGRGGFQDITIPWTRQNGRNFRFSEPLVIAV

PARAQIGRHSQGLMTSIIYIAVQDTRFIDVMDORTRELEPFGFELAVACA

TSVWCAVCPKVNTHAGCHRQLTLEIKEQGKFWMSDWTHATEDVLAQDDEMG

VEVRDEGSLSRGVYGEALKKAKTREGTPESSALASVRVILTFERELDSTKPPR

ERDVGAGTRIAQEVASGAVILRUEGGVPLDRAQDIAVIGPSAQOPKVWGLGSSY

VERDFANAPLUTITDVGSGSRVGSVGEETGKAPVPAFVCEVTPPSGGY

TYDGRILKUYDGLYLAIRRGGNSLQDIDGAPITGPGVJPSVWLFKGENTI

OMTGAAVPGGSLDVLWTPGHAOREDAVVERADSVDVAVFVADGAETATTA

LSPGTDOKLIDAVASVNPWVWVNTGSSWVPLDKRAVILWVWYGOARAA

LIFGDABPGGRUTOTPEVSDRTPVGGDPAFPGVGDKWHISCEISYRWDREGVD

PIEPFGHGLSLSTTEETDPUVERTRQDLYTVTVENTGGRGSIVVQYLGPPFQVPL

DOAPRQQLAGYQKVVELARGETKRVVRAELQHDEAGGWKLGGKRAVEIGSSR

DIDLRADINL"

gene

/gene="eryB1II"

CDS /codon\_start=1

/transl\_table=11

/protein\_id="CAA74703.1"

/db\_xref="SPREMBL:000463"

/db\_xref="MILVLOGKCRGNN"

/translation="MILVLOGKCRGNN"

BASE COUNT 383 a 966 c 982 g 429 t

ORIGIN

Query Match 7.1%; Score 178.2; DB 1; Length 2760;

Best Local Similarity 54.0%; Pred. No. 1.2e-12;

Matches 416; Conservative 0; Mismatches 343; Indels 12; Gaps 2;

QY 1591 GCGCACGGCTCCCTCGCGTGGGGCTGCAAGGTCATGACGACGGCAATCGAA 1650

Db 978 GCGGGCGGTTCTGCACGTCAGTCAGCTGGTCACCCGGCAGCCACGGGAG 919

QY 1651 AACGCTCTGCCTCGCAAGGAGACGACAGTCATCTCGCGGCTAACGCC 1710

Db 918 TTGACGCGGGTGGAGCGCGGCGACGCCAGCTGAGTCGGTGTTCGCTAACGCC 859

QY 1711 GACTGGGAGCGAGGGCGACCCGGGACGACGACGGTCATCTCGCGGCTTGAGCAC 1770

Db 801 CTGATCAGCCCTCGCTCGTCACCCGACACCGTGTGGTGTCAACGCCAGCG 742

QY 1831 CCCGAGGATGTCCTGCTGAGCCACGCGCCACGTCGAGTCAGGCTGTGACGGGCC 1890

Db 741 TCGTCACGATGGCTGGTGGACAAGACCCGGGTCGACATGTTGTTACCGGA 682

QY 1891 AACGAGAGGGCACTCCATGCGAGTCGCTGCTTGACTACACCCCTCGGCAAG 1950

Db 681 CAGCGCGGGGAGGCAACAGGGCTGCTGGCGAGCCGGCGGGCGGG 622

QY 1951 CTGCCCCCTGAGCTTCCCAGGAAACCCGGCTCTGAGACA 2001

Db 621 CTCACCCAGACCTTCCCGTGTGCAAGGACGCCGGCGACCGGGCGC 562

QY 2002 TTCCGACCCAGGGGGGGCACCTGTCAGSGAGGAGCTGCTACGTTGCTACAGGTAC 2061

Db 561 TTCCCGGGCTGGAGGAGGTGCACTCTCGAGGGCTCTCTGGGTACCGCTGG 502

Db 441 TTGGAGGGCACCGACCCGGTGTGAGCCACCGCGACGGCTCGACGGTACGGTAC 382

QY 2062 TAGACTTTCGCGACAAGGAGCTCAATTCCCTTSGCCAGGCCUGTCACACT 2121

Db 501 TAGACCGGGAGGGCTGGCCCTTTCGGTGTGGCACCGGGCTTCACACCG 442

QY 2122 TTGGCTTCTTCAACTCTCGTGTCTACAGAGGCCAACGTGAGGTGRCCTCTCC 2181

Db 381 GTGGCCACACCGGTGAGCCGGCCAGGGCTGTCAGCCACCGCGACGGCTCGACGGTACGGTAC 322

QY 2242 CAGCGCCCAAGATTACCCCGGCTCAAGAGCTCAAGGCTTCGCAAGGCTGAACTG 2301

Db 321 CGCGAGGTTCCGCTGGACCAAGGACCCGGCAGCTGGCCGGTACCAAGGGCGAAC 262

QY 2302 CAGCCCCGGAGGACGAGGGCGCTGACAATGAGGAGGAGGAGTAC 2352

Db 261 GCACCGGGAGGACCAAGGCGCCGGTCACTGCGGAGGGCCCTC 211

Search completed: May 4, 2003, 03:05:15

Job time : 4797.77 secs

